AD_____

Award Number: W81XWH-06-1-0449

TITLE: Early Detection of Breast Cancer via Multi-plane Correlation Breast Imaging

PRINCIPAL INVESTIGATOR: Amarpreet S. Chawla
Ehsan Samei, Ph.D.

CONTRACTING ORGANIZATION: Duke University Medical Center
Durham, NC 27710

REPORT DATE: April 2008

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1 Apr 2008	Annual Summary	31 Mar 2007 – 30 Mar 2008

4. TITLE AND SUBTITLE

Early Detection of Breast Cancer via Multi-plane Correlation Breast Imaging

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0449

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Amarpreet S. Chawl
Ehsan Samei, Ph.D.

E-Mail: asc14@duke.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Duke University Medical Center
Durham, NC 27710

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

One major deficiency of standard mammography is the camouflaging effect of overlapping structures in the projection images that limits the rendering of breast cancer. To minimize this effect, we proposed Multi-plane Correlation Imaging (MCI) technique. In this technique, multiple radiographic images of the breast are obtained from different angles. Angular information is used to identify and positively reinforce the lesion signals between different projections. In this phase of the study, we conducted a feasibility study of MCI using a mathematical model-based framework. Clinical utility of MCI was tested using a task that modeled clinical practice on clinically-realistic backgrounds of mastectomy specimens. The technique was then optimized to obtain maximum diagnostic information. Furthermore, a CADe based operator that takes advantage of the MCI configuration was developed. Results revealed that the peak performance for MCI at dose levels of one- and two-view mammography was achieved at 15 – 17 projections spanning an angular arc of ~45o, the widest angle tested in this study. Performance of optimized MCI exceeded that of standard mammography by 18%. Overall, MCI may prove to be potentially more accurate, and cost-and dose-effective.

15. SUBJECT TERMS
Mammography, Tomosynthesis, Multi-projection Imaging, Laguerre-Gauss Channelized Hotelling Observer, ROC, Detectability Index, SKE, Computer Aided Detection (CADe).

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	43	

Table of Contents

	Page
Introduction…….……..	**4**
Body….……..……..	**5**
Key Research Accomplishments…….……..…….	**17**
Reportable Outcomes…….…….	**17**
Conclusion…….	**18**
References……..	**19**
Appendix..……..…….	**20**

INTRODUCTION

Currently, mammography is the de facto technology for breast cancer screening. Although the technology is proven to be most sensitive among other comparable screening techniques,[1] it is not without limitation. The specificity of mammography is still relatively low; only 15-30% of suspected breast lesions recommended for biopsy are actually malignant.[2, 3] At the same time the false negative rate of mammography in dense breasts can be as high as 25%.[4]

In mammography, a 3D volume of anatomical structures is collapsed into a 2D image plane. The resultant image, therefore, is a consequence of projection of overlapping anatomical structures into a complex 2D image. As a result, any abnormality in the form of breast cancer may easily get hidden behind projections of a normal tissue structures, resulting in low sensitivity. At the same time, the overlapping structures may also mimic the pathology that the radiologists are looking for, leading to high rate of false positives. An imaging technique which may alleviate the limiting factor of overlapping anatomical structures and at the same time take advantage of standard mammography imaging technique, may therefore prove to be highly effective in breast cancer screening.

In this study, we are investigating the feasibility of a new imaging technique, namely, **Multi-plane Correlation Imaging (MCI)** in which a plurality of digital radiographic images of the breast are acquired within a short interval of time from slightly different angles. These images are similar to projection images acquired in standard digital mammography, except that each of the angular projections is acquired with lower dose level than used in standard mammography. These images are then processed by a computer algorithm which utilizes spatial correlation information between different angular projections to identify and positively reinforce the lesion signals between different projections, thus minimizing the fundamental limiting factor imposed by anatomical noise on detection of lesions.

For the first stage of this study, we investigated the optimum geometry of acquisitions in MCI in terms of the number of acquisitions and the angular span of those acquisitions that yield maximum performance in a clinical detection task. A simulated mass was embedded in the multiple projections to emulate a realistic clinical task. An approach based on mathematical observer model was developed to assess the detectability of the mass. Detectability was measured in terms of Receiver Operating Characteristics (ROC). The methodology first combined cues on the detectability of mass available from multiple projections into a combined ROC as final figure of merit to measure the system's performance. Next, a specific combination of number of angular acquisitions and the angular span of these acquisitions was determined that maximized the Area under the combined ROC curve (AUC). The combination that maximizes the AUC was deemed the optimized geometry for MCI acquisitions.

BODY

Specific Aim 1: Determine the set of acquisition parameters for an MCI study. (Months 1-9)

Task 1.1: Acquire multi-projection images of 10 cadaver breast specimens with and without lesions inserted at various settings of acquisition parameters. (Months 1-3)

Progress: Work for this task has been accomplished. The goal of this task was to acquire images that approximate mammographic backgrounds which could then be used for analysis in developing methodologies for the other specific aims of this study.

The first couple of months since the start of this grant were spent in securing an approval from the IRB office at Duke for retrospect use of the already acquired human subject data. The specific components of this study were determined to be in compliance with all applicable "HIPAA" regulations. Final IRB approval was received on 07/24/06.

Five mastectomy specimens were subsequently obtained from the pathology laboratory at the Duke hospital and imaged for this study. The specimens were acquired of subjects with proven history of malignant breast tumors, although no residual tumor was identified in the specimen. Multi-projection images of these specimens were then acquired per an approved IRB protocol. Under this protocol, images were acquired about the CC orientation from twenty-five different but fixed angular positions using a prototype clinical tomosynthesis system (Mammomat Novation^{TOMO}, Siemens Medical Solutions, Erlangen, Germany).[5] This system is also being used to conduct an ongoing tomosynthesis clinical trial.[6] The mastectomy specimens were immobilized by the compression paddle of the system. The compression values were in the clinical realistic 30 – 60 mm range. In order to increase the total sample size of our experiment using the available specimens, three of the five specimens were moved and compressed and imaged again, thus potentially imaging different and independent sets of anatomical configuration each time. Thus starting with five specimens, recompression resulted in image sets from four more specimens resulting effectively in an overall dataset of nine specimens. Acquisition of an additional specimen to meet the total sample size of 10 - originally proposed in this task - was deferred. This is because nine specimens acquired in this task were considered a sufficient sample size for robust statistical analysis undertaken in this study.

Task 1.2: Develop Laguerre-Gauss Channelized Hotelling Observer incorporating the spatial correlation information available via MCI technique. (Months 1-6)

Progress: The work on this task was started well in advance in anticipation of its application in two of the remaining aims.[7-9] As a result, this task was accomplished last year, as also noted in the last annual report. The goal of this task was to develop an observer model that acts like surrogate radiologists that could be used to prospectively investigate the expected clinical performance of MCI.

The observer model methodology developed as a part of this task served as an important foundation on which further theoretical analyses proposed in later tasks were built upon. These will be noted in the subsequent sections.

The outcome of this task was published in two articles the journal of Medical Physics in 2007 and 2008:

Chawla A., Samei E., Saunders R., Abbey C., Delong D., Effect of dose reduction on the detection of mammographic lesions: A mathematical observer model analysis, *Medical Physics* 34: 3385-3398, 2007.

Chawla A., Samei E., Saunders R., Lo J., Baker J., A mathematical model platform for optimizing a multi-projection breast imaging system, *Medical Physics* 35: 1337-1345, 2008.

Task 1.3: Apply the observer model on the image dataset to determine the optimum set of acquisition parameters, namely, maximum number of oblique views required, the specific angulations at which they should be acquired and the optimum dose level at which each of those should be acquired. (Months 7-9)

Progress: This task was the major focus of PI's efforts this year, and has now been accomplished. The specific methods and framework developed for this phase to meet the goals of this task will be detailed here.

I. Materials and Methods

a. Image database

Multiple 3 mm simulated masses were embedded at different locations within each of the 9 mastectomy specimen. Specifically, 84 non-overlapping 100 x 100 pixel regions of interests (ROIs) were identified for each projection image resulting in a total of 756 ROIs (84 ROIs*9 specimens) for each of the 25 angular projections. The 756 ROIs without the embedded masses were also used as signal-absent dataset for control purposes. Across projection images of each specimen, the projections of masses were added so to simulate 3D masses "implanted" inside the specimens at a distance of 3 cm above the detector surface.[10] Fig. 1 shows example images of a mastectomy specimen acquired at -22°, 0° (CC), and -23°, respectively.

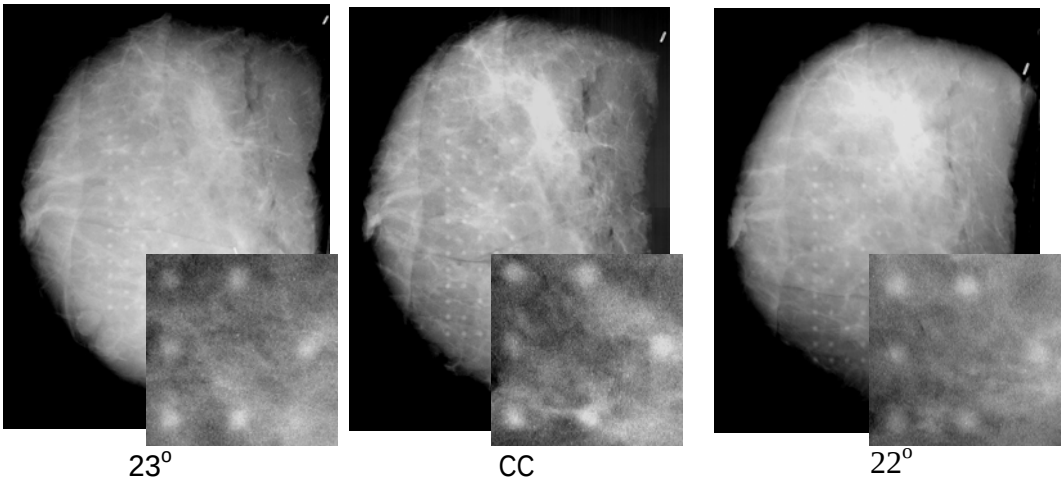
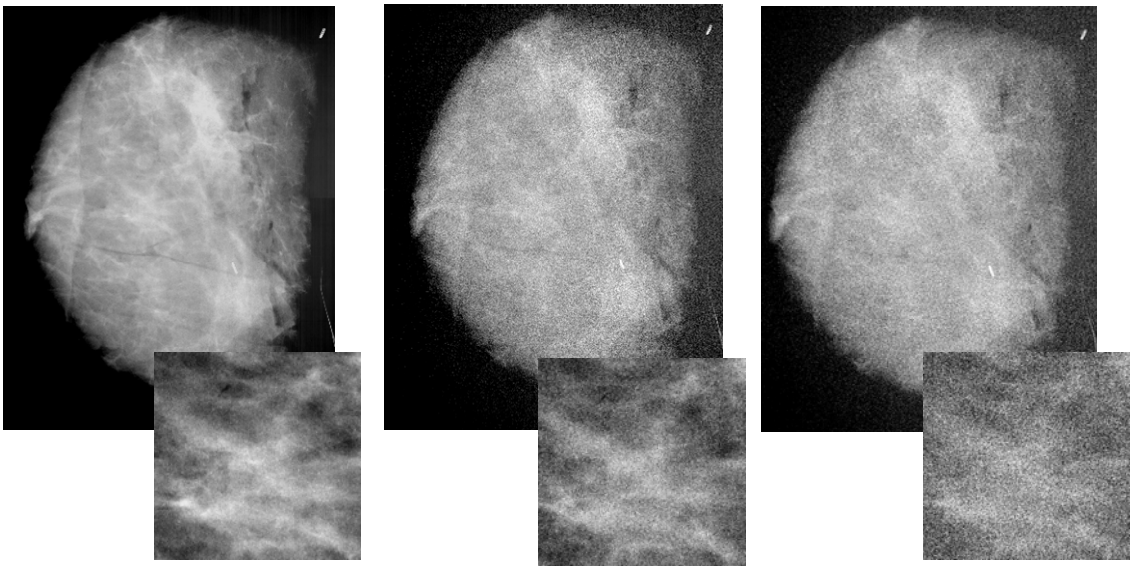


Fig. 1: Example projection images of mastectomy specimens with 3 mm simulated lesions embedded at the center of spatially distributed 84 ROIs. Note: contrast of the lesions was enhanced manifold for display purposes only.

Following lesion addition, a noise modification routine was used to add radiographic noise to each of the lesion-supplemented images to create images with a noise appearance similar to that caused from reduction in radiation dose. The routine was based on an algorithm reported earlier.[11] Several sets of dose-reduced images were simulated such that the cumulative dose of the 25 projections resulted in 10 discrete dose levels between 0.5D to 12.5D, where D is the typical single-view mammography dose level. Fig. 2 shows a representative CC projection of a specimen at three different noise/dose levels.



(a) (b) (c)

Fig. 2: Example images of a mastectomy specimen at clinical dose level (a), and with added noise corresponding to 50% and 25 % clinical dose level.

b. Optimization Framework

To optimize the acquisition scheme of MCI, the principal acquisition components, namely, dose, number of acquisitions and the angular range were systematically changed in each of the imaging modes to determine which one of the many possible combination of acquisition parameters maximized the detection of embedded lesions.

Optimization was investigated under two acquisition dose conditions based on the number of projections used, namely, iso-image dose conditions and iso-study dose conditions. Under iso-image dose condition, the dose level of each angular projection remained unchanged, thus resulting in increased dose level with an increase in the number of projections. Under the second dose condition, i.e., iso-study dose condition, the total dose was kept constant regardless of the number of projections used. Using the noise modified images, 3 clinically relevant iso- study dose conditions were simulated, namely, 1, 1.5, and 2 times single-view mammography.

For each dose level, the angular projections were systematically changed between 1 and 25 projections, while the angular ranges were varied in the 7.5^{o} – 44.8^{o} range. Finally, the performance of the system was determined as a function of the number of angular projections and the total angular span of those projections. A third-order polynomial fit was applied to fit the existing data points. The combination that yielded the maximum performance was deemed the optimized acquisition parameters set. Thus by changing the total dose level, a controlled yet comprehensive evaluation of optimization scheme in terms of dose, number of projections, and angular span was performed. Fig. 3 provides a visual illustration of this multi-factorial optimization scheme.

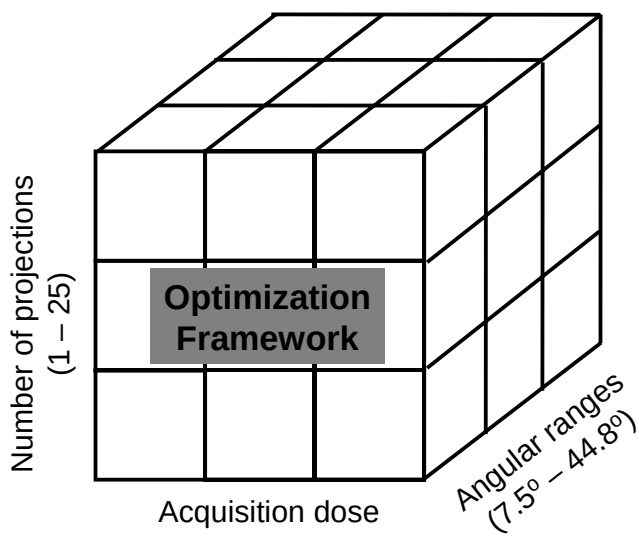


Fig. 3: Schematic of the optimization space used in this study to analyze MCI.

c. *Evaluation of Detection Performance*

The diagnostic performance of the system was evaluated in terms of the detectability of the embedded masses. Towards that a signal detection approach based on Laguerre-Gauss channelized Hotelling observer (LG CHO) developed as a part of previous task was used.[12] LG CHO takes into account the image and signal statistics to quantify the image quality and thus acts as a surrogate human observer. The detectability of the mass was measured in terms of Receiver Operating Characteristic (ROC) curves and the area under ROC curves (AUC).

For assessing the performance on MCI, the premise of CI, first the 756 lesion-supplemented ROIs per angle were extracted. Next, based on the known characteristic of the embedded lesion, the detectability of the lesions was determined using the LG CHO observer model approach. Detectability was measured at each of the 25 angular projections thus resulting in 25 ROC curves. ROCs corresponding to a given acquisition configurations were then combined using a decision fusion technique based on Bayesian statistics.[10, 13] The area under the combined ROC curve (AUC) was employed as an overall figure of merit representative of clinical performance of the projection images at that acquisition configuration.

II. Results from Task 1.3

Fig. 4 shows variation in AUC with the number of angular projections at different angular ranges for MCI, under iso-image dose condition (dose level of each projection was fixed at half that of typical single-view mammography). Regardless of the angular span, the AUCs first increased with the increase in the number of projections but leveled off beyond a certain number of projections. The maximum value of AUC, however, increased with an increase in the angular span. Most importantly, the peak performance was between 15 and 20 projections for an angular span of about 45°.

Figs. 5 and 6 shows variation of AUC with number of projections and angular spans under fixed dose levels (iso-study dose condition) at total dose levels equivalent to those for two and single-view mammography, respectively. There was an improvement in diagnostic performance when information from multiple images was combined, confirming the benefit of MCI over standard mammography. However, for all angular spans, the AUC first increased and then decreased as the number of projections was increased. The number of projections at which the AUC values peak was dependent on the angular span. Most noteworthy, the maximum AUC value was obtained at an angular span of 44.8^{o} with 15 – 17 projections.

While the total dose in Figs. 5 and 6 were fixed, Figs. 7 and 8 shows variation of AUC at three different dose levels, of 1, 1.5, and 2 times that of single-view mammography. Two different angular spans of 7.5^{o} (Fig. 7) and 44.8^{o} (Fig. 8) are shown. At each dose level, performance is optimized at a particular number of projections. Regardless of the angular spans, the AUC values increased by increasing the dose level.

Figs. 5 and 6 suggest that the optimum number of projections is dependent on the total angular span used. Figs. 9 and 10 summarize that finding. Fig. 9 shows the number of projections that

yield maximum AUC at different angular ranges, while Fig. 10 shows the corresponding AUCs at each of those angular ranges. The maximum AUC is obtained using a 44.8^o angular span and 15 – 17 projections. As 44.8^o was the maximum angular span tested in this study, it is expected that a wider angular span might yield even higher performance. The slope of the linear fit in Fig.9 reveals that for the optimum angular separation that realizes maximum performance in MCI is approximately 2.75^o.

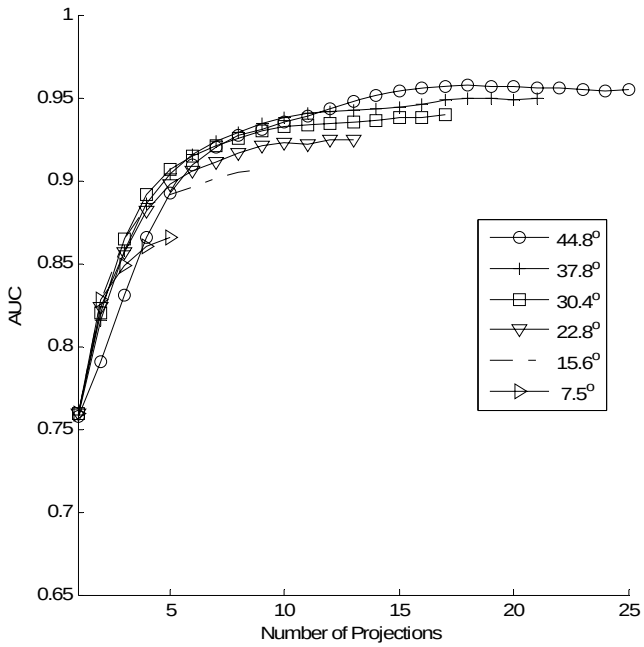


Fig. 4: Variation of AUC with number of projections for MCI at different acquisition dose levels under *iso-image dose condition*. The angular span of these projections were in the 7.5^o – 45° range.

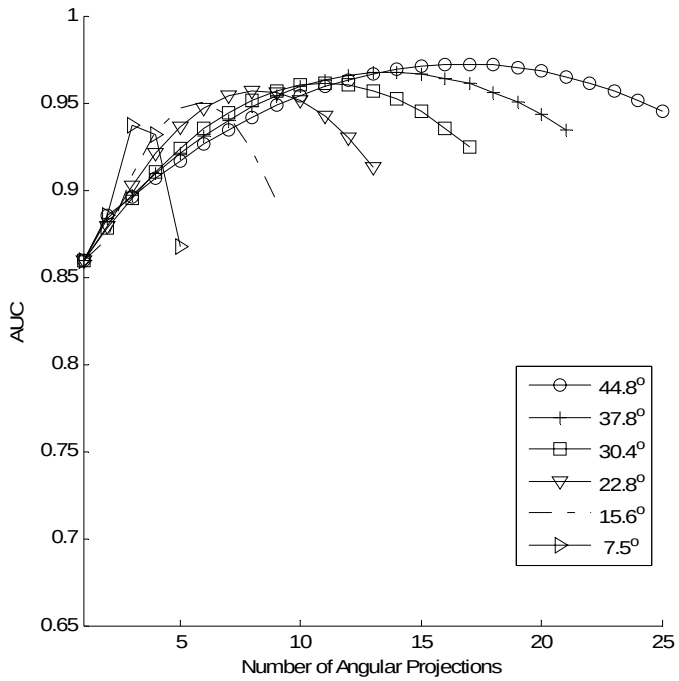


Fig. 5: Variation of AUC with the number of projections for MCI under iso-study dose conditions at different angular ranges in 7.5° - 45° range. The total dose level was fixed to that of two-view mammography.

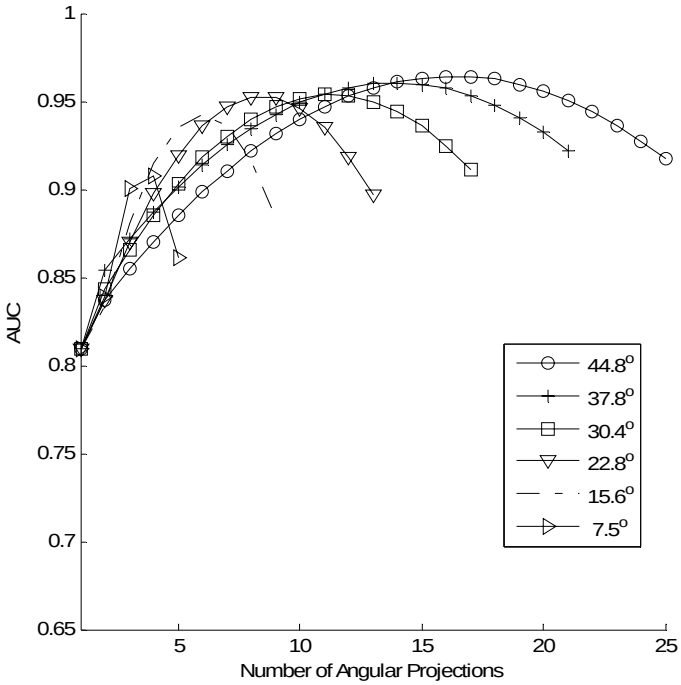


Fig. 6: Variation of AUC with the number of projections for MCI under iso-study dose conditions at different angular ranges in 7.5° - 45° range. The total dose level was fixed to that of two-view mammography.

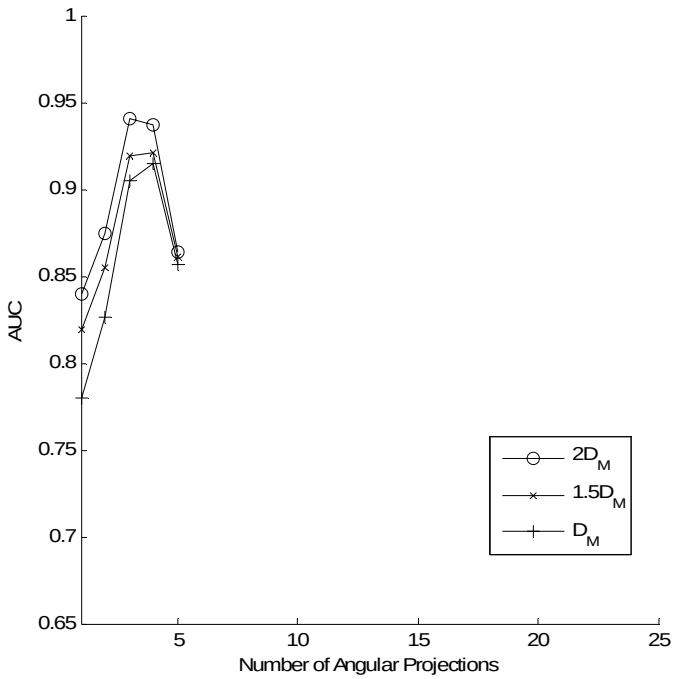


Fig. 7: Variation of AUC with the number of projections for MCI under iso-study dose conditions at an angular range of 7.5^{o}. These values are also plotted at different dose levels (denoted in the legend as the multiples of that of typical single-view mammography)

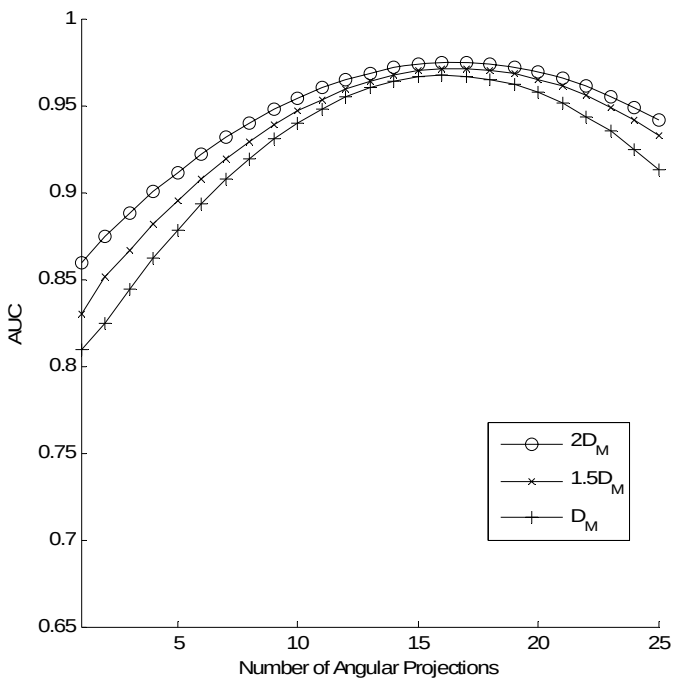


Fig. 8: Variation of AUC with number of projections for MCI under iso-study dose conditions at angular ranges 44.8^{o}. These values are also plotted at different dose levels (denoted in the legend as the multiple of that of typical single-view mammography)

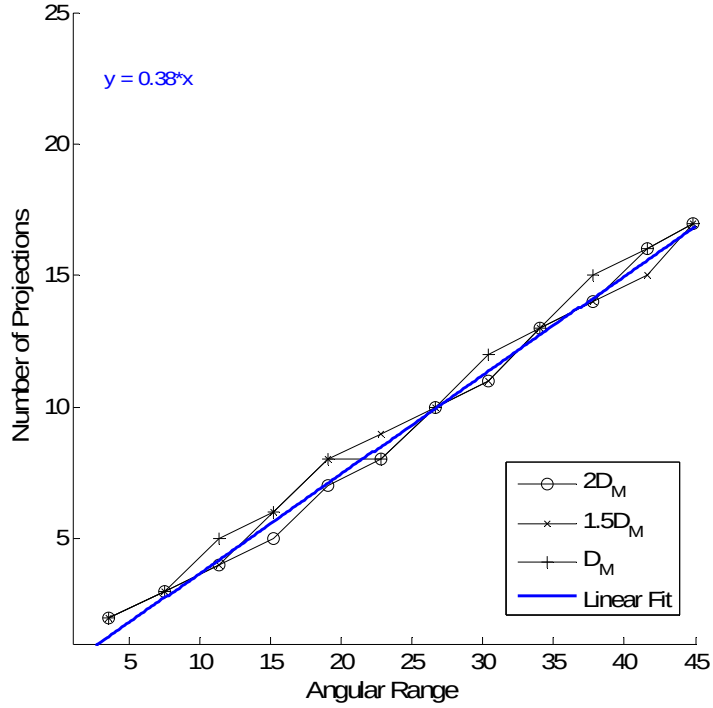


Fig. 9: The number of projections per angular range that yield maximum AUCs for MCI. These are plotted for different dose levels (denoted n the legends as the multiples of that of single-view mammography).

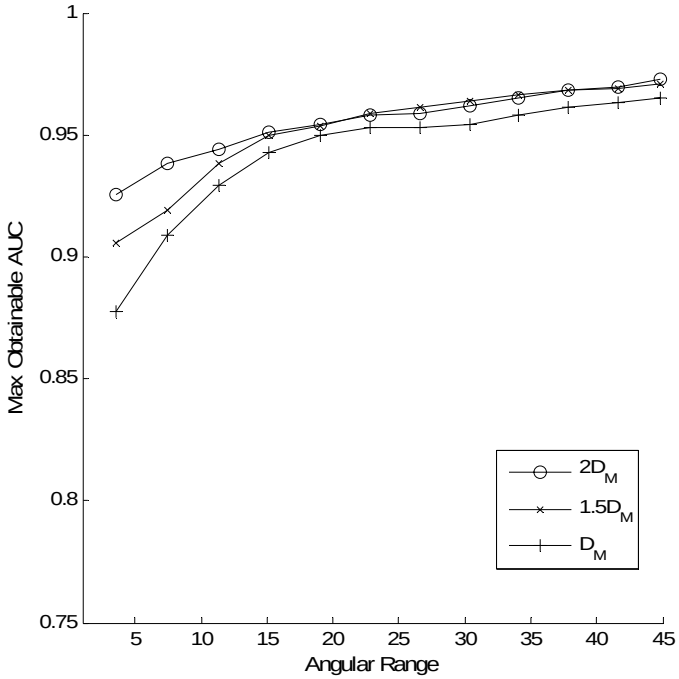


Fig. 10: Maximum obtainable AUC values for different angular ranges for MCI. These are plotted for different dose levels (denoted n the legends as the multiple of that of single-view mammography).

III. Conclusions of Task 1.3

Sub-optimized implementation of MCI can potentially compromise its maximum achievable diagnostic performance. In this task, we developed an algorithmic observer-based framework to assess the impact of various acquisition parameters of MCI performance, namely, dose, the number of projections, and the angular span. The study demonstrated the interplay of anatomical and quantum noise in the overall performance. The following conclusions may be drawn from this study:

1) Increasing the number of projections while keeping the overall dose and angular span constant decreased the performance of MCI.

2) Increasing the angular span and acquisition dose level improved the maximum obtainable AUC.

3) The number of projections required to maximize performance was found to be linearly related to the angular span. This number was found to be independent of the acquisition dose level. The best clinical performance was obtained when the angular separation between each projection was approximately 2.75^{o}.

4) Finally, the results revealed that the peak performance for MCI at the clinically relevant dose levels of one- and two-view mammography was achieved at 15 – 17 projections spanning an angular arc of $\sim 45^{o}$, the widest angle tested in this study.

Work for this specific task will be submitted for publication in the journal of *Medical Physics*.

Specific Aim 2: Extend single-view CAD processing methods used for conventional mammography for MCI implementation using a multi-plane correlation rule. (Months 10-21)

Progress: Work for this task has been completed as well. A CADe processor developed earlier for standard projection technique was extended to take advantage of the MCI configuration.[14] Specifically, 25 angular projections/case available in MCI were first filtered using a modified adaptive elliptical gradient convergence filter creating a blurry estimate of the anatomical background and highlighting suspicious abnormalities in the images. Following filtration, the suspicious regions were segmented with a grayscale duration technique [15]. The segmentation was optimized to highlight structures with sizes similar to the expected 3 mm embedded lesion. The segmented suspicious regions were analyzed for nine morphological features. These features were combined using a Bayesian decision fusion scheme to reduce false-positives [16]. The false-positive reduction program was trained using a genetic algorithm to find optimum feature thresholds that eliminated the greatest number of false positives while maintaining high sensitivity. The result was a set of 25 binary images per case, each showing potential locations of the embedded mass.

Next, the binary images per case obtained by the single-view CADe routine were processed to incorporate the MCI configuration. Specifically, a shift and add reconstruction technique was applied to generate a stack of 20 image slices resulting in a CADe-enhanced volume of image slices within which the potential lesion was segmented. The stack of slices was then collapsed (summed) into a single 2D image comparable with the central (CC) projection and containing

information from all the processed projections. The collapsed 2D image brought into focus the most suspicious regions, while the regions with a less likelihood of a presence of lesion were blurred out. Finally, a thresholding mechanism was applied to pick the region with the suspected pathology, thus providing a 2D contour map of the possible locations of the lesion.

To evaluate the performance of MCI, the 2D contour map was compared to a truth file. The truth file was defined as a binary mask of area that encompasses the known locations of the embedded lesion taking into account its spatial displacement across all the 25 projections. If a region on the 2D contour map overlapped the true lesion area, a true-positive finding was registered. All other regions that did not overlap were counted as false-positive findings. Using this rule, free-response receiver operating characteristics (FROC) curves were generated.

Fig. 11 shows the result of the CAD processor. Shown in the figure is a representative case with true location of an embedded lesion and the true positive and false positives findings of the CADe processor projected on the CC image. For reference, the angular projections of -22.3^{o} (a), 0° - CC orientation (b), and 23.1^{o} (c) that were used to boost the performance of CAD are also shown.

A new CADe processor was developed for multi-projection Correlation Imaging (CI) that takes advantage of the geometrical correlation information accrued from the available multiple projections to improve specificity of the CI system. The performance of CADe processor was found to be robust as it is successfully able to locate the suspected lesions.

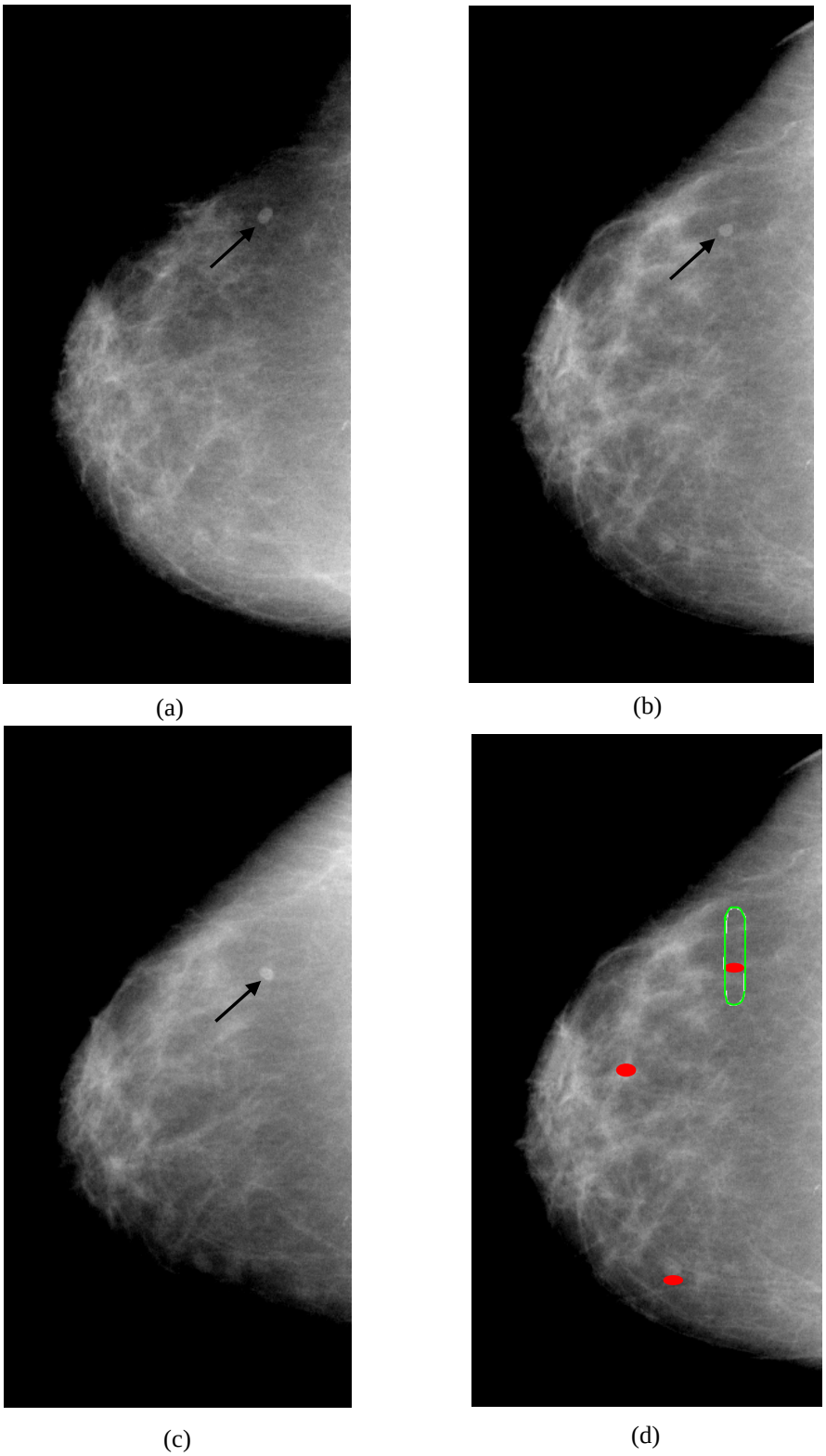


Fig. 11: (a), (b), (c) show projection images of a breast acquired by the multi-projection system at –22.3^{o}, 0° (CC orientation), and 23.1^{o}. The arrows show the locations of the embedded 3D lesion at these projections. (d) shows the CC projection image with suspected locations of lesion marked by the CADe processor in red. The location of the true lesion is encompassed in the green mark. The locations where the red regions intersect the green mark are noted as true-positive findings. (Note: the contrast of the lesions was enhanced manifold for display purposes only.)

The work resulting from this task will be submitted to the journal of Academic Radiology.

Specific Aim 3: *Fine-tune the optimized formalism of step 1 for CAD algorithm and evaluate its performance.*

The work for this task is under progress. We have, however, initiated the task of integrating the CADe processor with the observer model methodology. Since the work is in its preliminary stage, we will defer a report on this task to a subsequent annual summary.

KEY RESEARCH ACCOMPLISHMENTS IN YEAR 2007-08

Two of the four specific aims outlined in this grant were successfully met. The work for specific aim 3 was initiated. The methodologies developed in this phase of the project are key to meeting the goals of the remaining specific aims of this study. Specifically, the following were accomplished:

- The observer model methodology developed earlier was extended to incorporate the MCI configuration to quantify the available diagnostic information.
- Mastectomy specimens were procured and imaged to test our theoretical model.
- Building on the model, an optimization framework was developed to investigate the best combination of acquisition parameters in MCI that will maximize its clinical utility.
- A robust CADe processor for MCI was developed. It will be used to meet the goals of specific aim 3, i.e. to substantiate and fine-tune the optimization results of observer model.
- Optimized MCI was demonstrated to potentially improve overall accuracy over standard mammography.

REPORTABLE OUTCOMES

This work resulted in the following journal articles and conference proceedings. The names of the fellow (**Chawla**) and mentor (**Samei**) are boldfaced for emphasis.

Refereed Journal Publications:

1. **Chawla A.**, **Samei E.**, Saunders R., Lo J., Baker J., A mathematical model platform for optimizing a multi-projection breast imaging system, *Medical Physics* 35: 1337-1345, 2008. (Attached in the Appendix).
2. **Chawla A.**, **Samei E.**, Saunders R., Abbey C., Delong D., Effect of dose reduction on the detection of mammographic lesions: A mathematical observer model analysis, *Medical Physics* 34: 3385-3398, 2007. (Attached in the Appendix).

Conference Proceedings:

3. **Chawla A.**, **Samei E**., Lo J.Y., Mertelmeier T., Multi-projection Correlation Imaging as a new Diagnostic Tool for Improved Breast Cancer Detection, *Proc. IWDM IX*: 635-642, 2008.
4. **Chawla A., Samei E.**, Saunders R.S., Lo J.Y., and Singh S., Optimized acquisition scheme for multi-projection correlation imaging of breast cancer, *Proc. SPIE Medical Imaging* 6915, 691528: 1-8, 2008.
5. Singh S., Tourassi G.D., **Chawla A.**, Saunders R.S., **Samei E**., Lo J.Y., Computer-aided detection of breast masses in tomosynthesis reconstructed volumes using information-theoretic similarity measures**,** *Proc. SPIE Medical Imaging* 6915**,** 691505:1-8, 2008.
6. **Chawla A**., Samei E., and Abbey C., A mathematical model approach toward combining information from multiple image projections of the same patient, *Proc. SPIE Medical Imaging* 6510(1K): 1-11, 2007.
7. **Chawla A.,** Saunders R., Abbey C., Delong D., **Samei E**., Analyzing the effect of dose reduction on the detection of mammographic lesions using mathematical observer models, *Proc. SPIE Medical Imaging* 6146(0I): 1-12, 2006.

CONCLUSIONS

We have demonstrated feasibility of multi-plane correlation imaging (MCI) as a technique for improved breast cancer detection. An optimization framework was developed to maximize the diagnostic performance of MCI. It was revealed that the peak performance for MCI at the clinically relevant dose levels of one- and two-view mammography may be achieved at 15 – 17 projections spanning an angular arc of ~45°, the widest angle tested in this study.

Overall, compared to mammography, MCI was demonstrated to be potentially more accurate, and cost- and dose- effective. Future work with CAD will substantiate the results of the present findings.

REFERENCES

[1] L. Tabar and P. B. Dean, "The Control of Breast Cancer through Mammography Screening," Radiologic Clinics of North America **25**, 993-1005 (1987).

[2] D. B. Kopans, "The positive predictive value of mammography," American Journal of Roentgenology **158**, 521-526 (1991).

[3] D. D. Adler and M. A. Helvie, "Mammographic biopsy recommendations," Current Opinion in Radiology **4**, 123-129 (1992).

[4] R. E. Bird, T. W. Wallace and B. C. Yankaskas, "Analysis of cancers missed at screening mammography," Radiology **184**, 613–617 (1992).

[5] S. Singh, G. D. Tourassi and J. Y. Lo, "Breast mass detection in tomosynthesis projection images using information-theoretic similarity measures," Proc. SPIE **6514**, 651415 (2007).

[6] M. P. Eckstein, C. K. Abbey and E. O. Bochud, "The effect of image compression in model and human performance," Proc. SPIE **3663**, 243-252 (1999).

[7] M. P. Eckstein, C. K. Abbey and J. S. Whiting, "Human vs. model observers in anatomic backgrounds," Proc. SPIE **3340**, 16-26 (1998).

[8] A. E. Burgess, F. L. Jacobson and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise," Med. Phys. **28**, 419-437 (April 2001).

[9] R. S. Saunders, E. Samei and C. Hoeschen, "Impact of resolution and noise characteristics of digital radiographic detectors on the detectability of lung nodules," Med. Phys. **31**, 1603-1613 (2004).

[10] E. Samei, M. Flynn and W. Eyler, "Simulation of subtle lung nodules in projection chest radiography," Radiology **202**, 117-224 (1997).

[11] A. E. Burgess, F. L. Jacobson and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise," Med. Phys. **28**, (2001).

[12] J. L. Jesneck, L. W. Nolte, J. A. Baker, C. E. Floyd and J. Y. Lo, "Optimized approach to decision fusion of heterogeneous data for breast cancer diagnosis," Med. Phys. **33**, (2006).

[13] R. S. Saunders and E. Samei, "A method for modifying the image quality parameters of digital radiographic images," Med. Phys. **30**, 3006-3017 (2003).

[14] R. S. Saunders, E. Samei, N. Majdi-Nasab and J. Y. Lo, "Initial human subject results for breast bi-plane correlation imaging technique," Proc. SPIE **6514, 651423**, 1-7 (2007).

APPENDIX

1. **Chawla A.**, Samei E., Saunders R., Lo J., Baker J., A mathematical model platform for optimizing a multi-projection breast imaging system, *Medical Physics* 35: 1337-1345, 2008……21

2. **Chawla A.**, Samei E., Saunders R., Abbey C., Delong D., Effect of dose reduction on the detection of mammographic lesions: A mathematical observer model analysis, *Medical Physics* 34: 3385-3398, 2007…….....30

A mathematical model platform for optimizing a multiprojection breast imaging system

Amarpreet S. Chawla[a)]
Duke Advanced Imaging Laboratories, Departments of Radiology and Biomedical Engineering, Duke University, Durham, North Carolina 27705

Ehsan Samei
Duke Advanced Imaging Laboratories, Departments of Radiology, Medical Physics, and Biomedical Engineering, Duke University, Durham, North Carolina 27705

Robert S. Saunders
Duke Advanced Imaging Laboratories, Departments of Radiology and Physics, Duke University, Durham, North Carolina 27705

Joseph Y. Lo
Duke Advanced Imaging Laboratories, Departments of Radiology, Medical Physics, and Biomedical Engineering, Duke University, Durham, North Carolina 27705

Jay A. Baker
Duke Advanced Imaging Laboratories, Department of Radiology, Duke University, Durham, North Carolina 27705

(Received 19 September 2007; revised 1 January 2008; accepted for publication 29 January 2008; published 12 March 2008)

Multiprojection imaging is a technique in which a plurality of digital radiographic images of the same patient are acquired within a short interval of time from slightly different angles. Information from each image is combined to determine the final diagnosis. Projection data are either reconstructed into slices as in the case of tomosynthesis or analyzed directly as in the case of multiprojection correlation imaging technique, thereby avoiding reconstruction artifacts. In this study, the authors investigated the optimum geometry of acquisitions of a multiprojection breast correlation imaging system in terms of the number of projections and their total angular span that yield maximum performance in a task that models clinical decision. Twenty-five angular projections of each breast from 82 human subjects in our breast tomosynthesis database were each supplemented with a simulated 3 mm mass. An approach based on Laguerre–Gauss channelized Hotelling observer was developed to assess the detectability of the mass in terms of receiver operating characteristic (ROC) curves. Two methodologies were developed to integrate results from individual projections into one combined ROC curve as the overall figure of merit. To optimize the acquisition geometry, different components of acquisitions were changed to investigate which one of the many possible configurations maximized the area under the combined ROC curve. Optimization was investigated under two acquisition dose conditions corresponding to a fixed total dose delivered to the patient and a variable dose condition, based on the number of projections used. In either case, the detectability was dependent on the number of projections used, the total angular span of those projections, and the acquisition dose level. In the first case, the detectability approximately followed a bell curve as a function of the number of projections with the maximum between 8 and 16 projections spanning angular arcs of about 23°–45°, respectively. In the second case, the detectability increased with the number of projections approaching an asymptote at 11–17 projections for an angular span of about 45°. These results indicate the inherent information content of the multiprojection image data reflecting the relative role of quantum and anatomical noise in multiprojection breast imaging. The optimization scheme presented here may be applied to any multiprojection imaging modalities and may be extended by including reconstruction in the case of digital breast tomosynthesis and breast computed tomography. © *2008 American Association of Physicists in Medicine*. [DOI: 10.1118/1.2885367]

Key words: optimization, acquisition parameters, ROC, AUC, Hotelling observer, LG CHO, multiprojection imaging, digital breast tomosynthesis, decision fusion

I. INTRODUCTION

Diagnostic radiology is increasingly embracing modalities that acquire multiple images of the same patient. The multiimage scheme alleviates the main shortcoming in standard projection imaging techniques: the overlap of anatomical structures (i.e., anatomical noise) that can partially or completely hide a pathology of interest.[1,2] This is done by har-

nessing the geometrical and statistical dependences between the multiple images available in a multiprojection system. One particular multiprojection acquisition modality, digital breast tomosynthesis, has indeed been shown to provide improved diagnostic performance as compared to the standard projection procedures.[3] However, the tomosynthesis technique is prone to reconstruction artifacts,[4,5] which might reduce the efficiency of tomosynthesis reading, even leading to higher false positive findings.[6] Therefore, there would be an advantage in using an imaging technique that could fuse information from multiple images, similar to tomosynthesis, but without the confounding effects of the reconstruction.

Toward that end, multiprojection correlation imaging (CI) was recently proposed as an adjunct technique to standard mammography or tomosynthesis.[7,8] This technique uses the unreconstructed angular projections acquired using an approach similar to tomosynthesis. The projection images are similar to standard mammograms, except that each image is acquired with a lower dose level than that in standard full field digital mammography. The information from the multiplicity of angular projections acquired is then combined to identify potential lesions. Practically, this combination can take different forms including scrolling the images manually or in cine mode, stereoscopic display of projections images, or computer-aided analysis of the multiple images. Thus, CI aims to augment the advantages of standard projection techniques with the proven benefits of multiprojection scheme, without the reconstruction artifacts that otherwise limit tomosynthesis, to deliver an improved diagnostic performance.[9–11]

While CI has notable potentials, in developing multiprojection CI, an important consideration is its data acquisition scheme. Multiple aspects of data acquisition can influence the performance of CI. Ideally, the data acquisition scheme should avoid patient motion, reduce patient discomfort, and maintain a total dose not exceeding that delivered in a standard projection technique. Beyond these, however, the diagnostic outcome of a multiprojection system is strongly dependent on the geometry of acquisition: the number of angular acquisitions and the total angular span of those acquisitions as geometry of acquisitions play a pivotal role in establishing correlation information in CI. It is therefore important that image acquisitions parameters be optimized to maximize the diagnostic information of such a system.

In this study, we investigated the optimal number and range of angular projections of a multiprojection breast CI system based on maximizing performance in a task that models clinical practice. The performance was measured in terms of detectability of an embedded simulated mass using a mathematical observer model.[12–16] As a key element of the study, three techniques were developed to combine the information content from individual angular projections to derive a combined receiver operating characteristic (ROC) curve that indicates the overall detectability of the mass. To optimize the geometry of acquisitions, the acquisition parameters were systematically changed to determine which one of the combination of parameters would maximize the area under the combined ROC. Optimization was investigated under two key dose considerations corresponding to a fixed total dose delivered to the patient and a variable dose condition in which the total patient dose increases with the number of projections acquired.

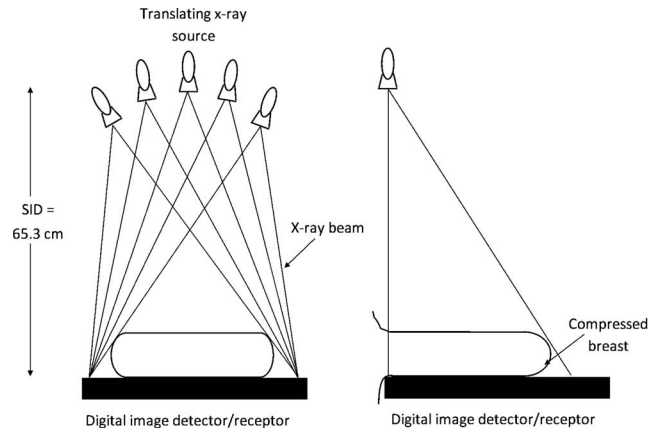


FIG. 1. The prototype multiprojection breast imaging instrument.

II. MATERIALS AND METHODS

II.A. Image database

The study employed a database of images of the left and right breasts of 82 subjects originally acquired as a part of our tomosynthesis clinical trial at Duke. Images were acquired about the CC or MLO orientation from 25 different but fixed angular positions by a prototype clinical multiprojection system, a modified Siemens' Mammomat Novation^{TOMO} (Fig. 1). The system used a selenium-based, flat-panel, digital mammography detector with an array size of 2816×3584 and a pixel pitch of 85 μm. The system used a tungsten target, a 50 μm rhodium filter, a source to image distance of 65.3 cm, and an isocentric gantry pivoting the x-ray tube about a point located 6 cm above the detector. The projection angles of the 25 images were varied within 45° angular range about the central orientation (CC or MLO) in steps of approximately 2°. The tube voltage ranged between 28 and 30 kVp to obtain consistent image contrast across different compressed breast thicknesses in the 3–8 cm range in our clinical trial. The total dose delivered to the patient from 25 angular acquisitions was equivalent to that delivered in a standard two-view mammographic screening procedure, with each angular projection at a dose level, D_θ, equal to 1/25th of the total clinical dose. All images were judged by a dedicated breast-imaging radiologist to be free from suspicious lesions.

492 regions-of-interest per projection angle (ROIs) (2 breasts $\times$ 82 patients $\times$ 3 ROIs/projection) of size 512×512 (43.5×43.5 mm^2) were extracted from the database. The displacement of the ROIs on the detector across the different angular projections was taken into account so that the ROIs from the same patient represented the same general volume of the breast. There was a slight difference in the breast volume being sampled as a function of the projection angle. However, this difference was found to be inconse-

quential since changing the ROI size was found to have a minimal effect on the overall performance of the observer model. From the pool of extracted ROIs, 264 were used for training the observer model, while the remaining 228 ROIs were used for testing.

II.B. Three-dimensional lesion smulation

A simulated mammographic lesion, 3 mm in diameter, was digitally inserted into the ROIs generated above. The size of the lesion enabled a difficult but clinically relevant detection task. The lesion was simulated in three dimensions (3D). To do so, first a two-dimensional projection profile based on a previously published model of lung and breast lesions was generated.[15,17,18] Starting from this profile, the surface of the central slice of the lesion was reconstructed using inverse radon transform, assuming that the lesion is isotropic and the different angular projections along the plane of the central slice would yield the same profile across the central slice. The central slice was then rotated about its diameter to complete the simulation of a three-dimensional (3D) lesion.

The 3D lesion was projected into the 25 angular projections assuming that the lesion was embedded at the center of the ROI volume (2–8 cm in compressed thickness) at a distance of 3 cm from the detector. The x and y coordinates of the projected lesion on the image plane were computed as

$$x_i = \frac{x(D + L\cos\phi) - zL\sin\phi}{D + L\cos\phi - z},$$

$$y_i = \frac{y(D + L\cos\phi)}{D + L\cos\phi - z}, \tag{1}$$

where (x,y,z) are the positional coordinates of any point on the lesion, (x_i,y_i) are the corresponding coordinates in the image plane, ϕ is the projection angle, and L and D are the distances of the pivotal point from the source and the detector, respectively. These equations were derived based on the trajectory of the acquisition system and are consistent with prior work.[19,20] Figure 2 shows projection of the 3D lesion on the detector from three different tube angular orientations of +22°, 0°, and −22° relative to the CC orientation.

The ratio of the contrast of the lesion to its diameter (set to 3 mm) was determined from published contrast/lesion thickness ratios based on the acquisition kVp, target/filtration combination, detector type, compressed breast thickness, and breast composition.[17] Since increased glandularity decreases lesion contrast, a 75/25% glandular /adipose breast composition was used to represent a difficult but clinically relevant detection task. The contrast ratio was further modified to take scattering into account. Toward that end, the scatter-to-primary ratio for the central projection was computed based on an earlier study.[17] This ratio was used as a representative value for all the projections, not reflecting slight variations in the scatter-to-primary ratio with angular projections.[21] Although consistent with the value reported in an earlier study, the variation of scatter to primary ratio with was ignored. The lesions were then added to the ROI in the log space such that the contrast of the lesion over the background was independent of the breast composition or thickness. The lesions were embedded onto the different ROIs to generate 492 signal-present ROIs for each angular projection. Figure 3 shows example images of ROIs with the embedded simulated lesion.

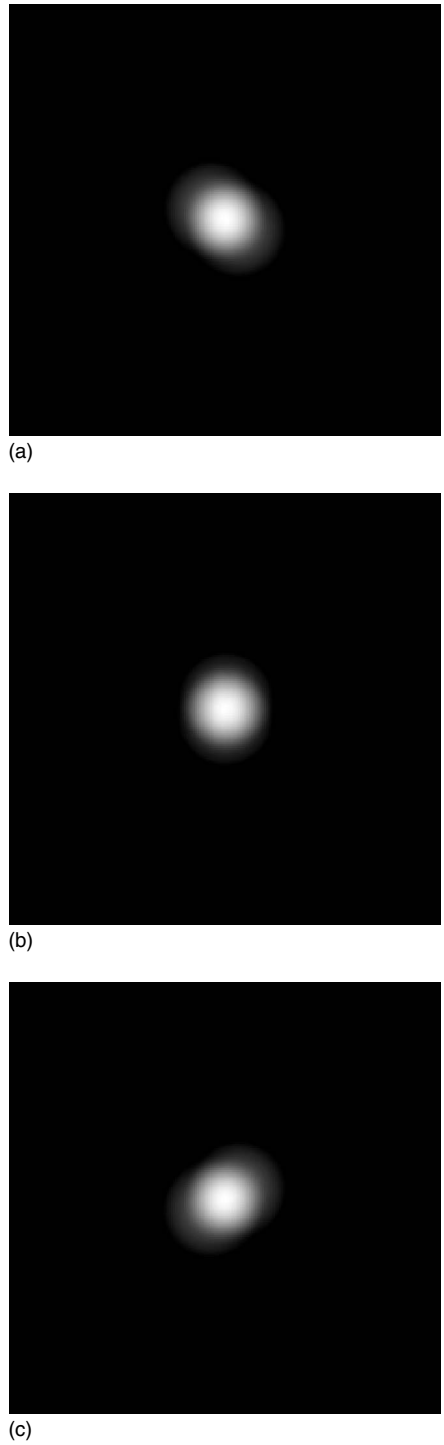


FIG. 2. Example images of projections of 3D model of a 3 mm simulated lesion assumed to be 3 cm above the detector. These lesions were embedded on tomographic projections to emulate the lesion-present mammographic background. (a), (b), and (c) show the projections with the tube orientation at +22°, 0°, and −22°, respectively, relative to the CC orientation.

II.C. Noise simulation

Following extraction of ROIs, a noise modification routine was used to add radiographic noise to each of the ROIs to create images with a noise appearance similar to that caused by reduction in radiation dose from D_θ. Noise equivalent quanta of a particular dose level was simulated using an algorithm reported earlier.[22] The algorithm accounted for the quantum noise variance, the detector transfer properties and its noise characteristics, and the impact of varying attenuation of breast structures. By changing the noise magnitude, 24 dose-reduction levels, corresponding to $D_\theta/20-D_\theta/25$, were simulated. These with the original clinically acquired images at D_θ resulted in images with 25 contiguously decreasing dose levels. Figure 3 shows examples of ROIs with different dose levels. While Fig. 3(a) shows the original clinically acquired image, Figs. 3(b) and 3(c) show the same ROI with added simulated noise corresponding to reduced dose level of $D_\theta/2$ and $D_\theta/25$, respectively.

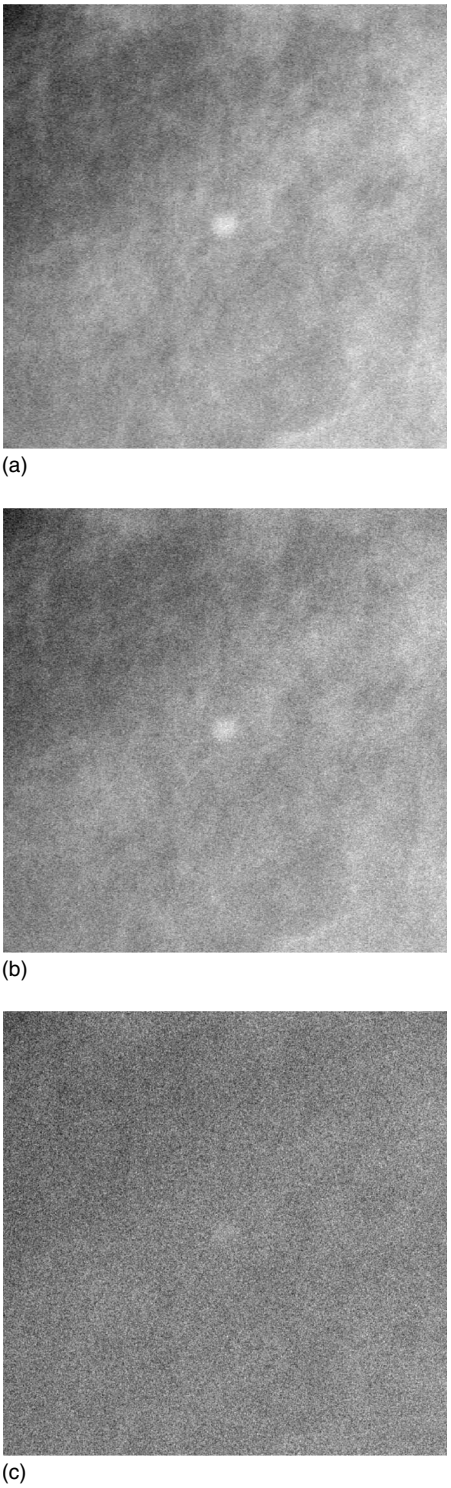


FIG. 3. Example projection images of ROIs with 3 mm simulated lesions embedded at the center. (a) shows the ROI of a clinically acquired projection with dose level, D_θ, equal to 1/25th that of standard mammographic screening. (b) and (c) show the same ROI with noise corresponding to 1/2 and 1/25th fraction of D_θ.

II.D. Mathematical observer model

Linear mathematical observer models, such as Hotelling observers, have been shown to predict human observer performance in clinically relevant visual tasks such as the detection of lesions in real anatomic backgrounds.[15,23–25] This study employed a variant of the Hotelling observer, namely the Laguerre–Gauss channelized Hotelling observer (LG CHO). LG CHO uses linear features that are product of Laguerre polynomials and Gaussians functions to reduce the dimensionality of the Hotelling observer, thus making the implementation mathematically tractable. The variance of the Gaussian function is related to the signal radius and is determined iteratively to maximize the area under the ROC curves (AUC).[15] A variance of 10 was used in this study. The present implementation further used a total of ten channels which is more than sufficient for characterizing isotropic signals used in our study (as shown in Fig. 2).[26]

The covariance matrix of the LG CHO was trained with signal-absent ROIs. Using a methodology previously published,[16] a set of signal-absent and signal-present decision variables were determined for each angular projection. Nonparametric ROC curves were then derived by simple thresholding on the probability density function (pdf) of the decision variables. While testing using the LG CHO, signal in each of the ROIs was analyzed with the signal-known exactly methodology, assuming that the lesions embedded in different ROIs within the same angular projection have approximately the same shape. The results were characterized in terms of the ROC for each of the 25 angular projections with the area under the mean ROC curve subsequently determined using the trapezoidal rule.

II.E. ROC fusion

Twenty-five ROCs obtained from the 25 angular projections are indicative of the performance of an observer as it analyzes each of those projections separately. Since the final clinical decision is based on a collective decision made from detectability cues gathered from each angular projection, it is essential to fuse the 25 ROCs into one final index of performance. Toward that end, two fusion methods were used and their performance compared to the average of the 25 ROCs.

In the first method, a weighted average of the signal-present decision variables from the 25 angular projections was computed to determine a final set of decision variables from which the combined ROC was derived. The weight assigned to the decision variable of an angular projection, θ, was based on the difference of detectability index between the value at the angular projection under consideration, d'_θ, and that at the central (CC or MLO) projection, d'_{CC}, and the absolute value of the angular separation between the two ($\theta-\mathrm{CC}$), as $w_\theta=(d'_\theta-d'_{\mathrm{CC}})/(|\theta-\mathrm{CC}|)$. The underlying hypothesis for this technique was that the overall detectability of the lesion should be dependent on its detectability at the oblique-angle projections and the proximity of those projections relative to the central orientation. While the overall detectability of a lesion should increase if its detectability at an oblique-angle projection is higher than at the central projection and vice versa, it should be higher still if the orientation of the oblique-angle projection is close to the central orientation causing its background to have a high correlation with that of the central projection (Thus, a smaller angular separation from the center provided a larger contribution of the correlation of the image information to the combined ROC.)

For the second ROC fusion method, a modification of a recently published Bayesian decision fusion algorithm was used.[27] In this method, first the ROC for each angular projection in the training data set of 264 ROIs was computed. For each of the 228 ROIs in the testing data set, a signal-present decision variable was then calculated by determining the response of the image embedded with ith lesion to ith lesion template (λ_1^i) and the corresponding signal-absent response by determining the response of the image itself (without the lesion embedded) to the ith lesion template (λ_0^i). Binary observer decision β_i to the ith image was computed as

$$\beta_i = \mathrm{step}(\lambda_1^i - \lambda_0^i) \Rightarrow \beta_i = \begin{cases} 1 & \text{if} \quad \lambda_1^i \geq \lambda_0^i \\ 0 & \text{if} \quad \lambda_1^i < \lambda_0^i \end{cases} . \tag{2}$$

The above equation implies that the threshold for correct observer outcome of an ith image is λ_0^i. The values of probability of true positive, p_d, and of false positives, p_f, corresponding to this threshold were then determined from the ROCs of the 25 angular projections in the training data set. Assuming that the binary decisions were statistically independent, the pdfs of the fused decision variables for signal-present and null hypothesis were then obtained as

$$P(\lambda_{\mathrm{fusion}}|H_1) = \prod_{\theta=1}^{25} (p_d^\theta)^{\beta_i}(1-p_d^\theta)^{1-\beta_i},$$

$$P(\lambda_{\mathrm{fusion}}|H_0) = \prod_{\theta=1}^{25} (p_f^\theta)^{\beta_i}(1-p_f^\theta)^{1-\beta_i}, \tag{3}$$

where θ is an index for angular projection. Having found the signal-present and signal-absent decision variables, the pdfs of each were computed from which the combined ROC curve and the AUC were deduced.

The second technique may be understood in the following way: given an image from angular projection, θ, assuming that the decision for signal present is 1, the probability of correct detection is p_d^θ. However, if the decision is 0, the probability that the signal is still present is the probability of false negative and, hence, $(1-p_d^\theta)$. Thus, the resultant probability of signal present in the image is $p_d^\theta\cdot(1-p_d^\theta)$. Assuming statistical independence of decisions among angular projections gives the pdfs of the fused decision variables of Eq. (3). In contrast to a genetic algorithm approach used in Ref. 25 to arrive at the binary decision, here we employed a simple thresholding approach that compares signal-present decision variable to a signal-absent decision variable to determine the binary decision.

II.F. Evaluating optimum acquisition parameters

For evaluating the effect of changing operating acquisition parameters on the performance of CI, different combinations of the number of angular projections and the angular spans were considered. A combined ROC was derived for each set of those acquisition parameters using the two ROC fusion methodologies. Finally, the area under each of the combined ROC curves (AUC) was determined as a function of the number of angular projections and the total angular span of those projections. A third-order polynomial fit was applied to fit the existing data points. The combination that yielded the maximum AUC was deemed the optimized acquisition parameters set.

The optimization was evaluated at two dose conditions:

(a) The isoimage-dose condition in which each angular projection considered for final figure of merit had the same dose level resulting in increasing total dose with each additional angular projection. Under this condition, the performance was evaluated by fusing decisions from N acquisitions. Ignoring the slight variations in dose levels with angular projections,[28] the N acquisitions, each at a dose level of D, resulted in a total delivered dose of ND.

(b) The isostudy-dose condition in which the total dose is divided among the projection images, such that incorporating additional angular acquisition would not result in increased total dose delivered to the patient. While evaluating performance from fusing decisions from N acquisitions, only acquisitions with dose levels of D/N were used, resulting in a fixed total delivered dose of D. Therefore, the optimum number of angular projections and the angular range determined in this case were independent of the dose delivered to the patient.

III. RESULTS

ROC curves and the corresponding AUCs obtained from the three techniques are shown in Fig. 4. The AUC value obtained from a single CC projection was 0.724, while that obtained from the average of 25 ROCs (corresponding to 25 projections) was 0.731. In comparison, the AUCs obtained

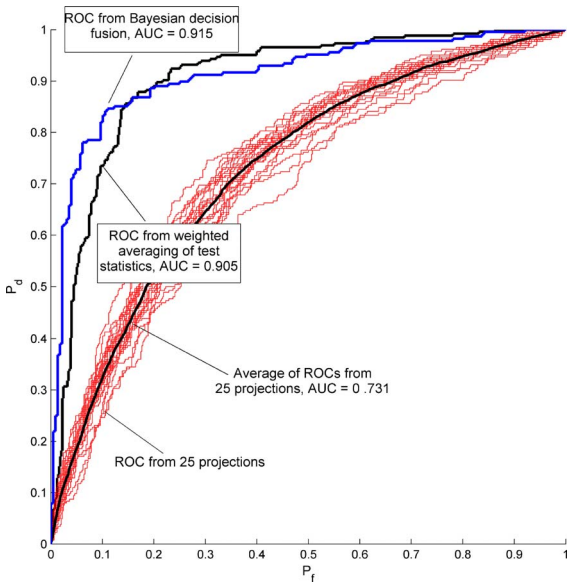


FIG. 4. ROCs of 25 projections obtained from a multiprojection imaging system and the average of those. Also shown are the ROCs obtained from the two fusion techniques. The angular span of the projections was 44.8°.

by combining information from 25 projections using the two combination techniques, namely weighted averaging of test statistics and the Bayesian decision fusion technique were 0.905 and 0.915, respectively. These results represent no modification to the dose level of individual projections.

Figure 5 shows variation of the AUC values as a function of the number of projections for different dose levels under isoimage-dose conditions. The total angular span of these projections was fixed at 44.8°. AUC for one projection was computed by averaging the AUCs across all the angular projections within this angular span. The dose level of each acquisition along a curve was constant and corresponded to the fraction of the clinical dose level indicated in the legend. While the AUC values monotonically increase with an increase in the number of projections at lower dose fractions of 1/11 and less, they appear to approach an asymptote with the increase in the number of projections for dose fractions of 1/7 and higher. Regardless, the AUC values increase with increasing dose levels.

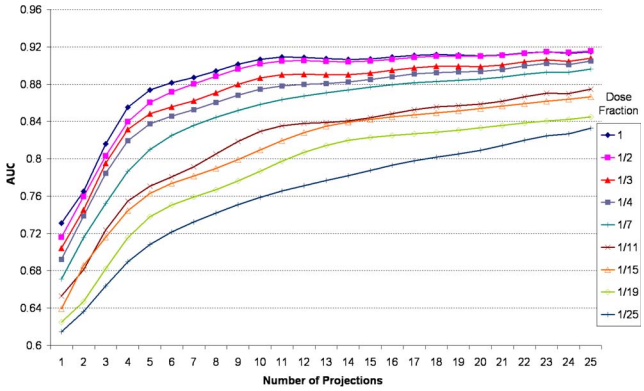


FIG. 5. Variation of AUC with a number of projections for different dose levels. Isoimage-dose condition was used implying that the dose level of each projection (D_θ) along a curve remains constant (i.e., more projections imply more dose to the patient). This dose level is indicated by the fraction of the clinical dose level in the legend. The Bayesian decision fusion technique was used for this analysis. The angular span of the projections was 44.8°.

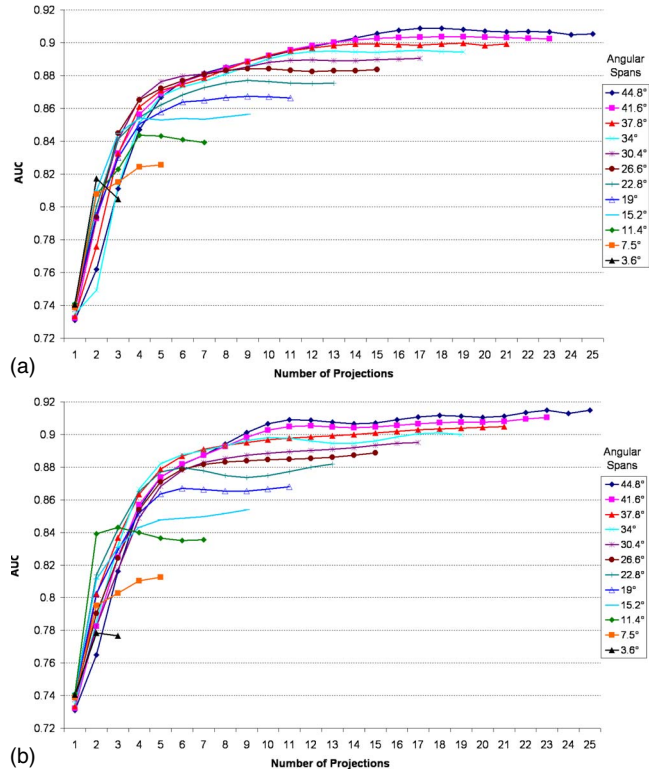


FIG. 6. Variation of AUC under isoimage-dose conditions for a different number of angular projections spanning a total angular arc in the 3.6° −44.8° range using (a) weighted averaging of test statistics techniques and (b) Bayesian decision fusion. The dose level of each acquisition was equal to 1/25th of the standard mammographic screening dose level leading to an increased dose level with an increasing number of angular projections considered to reach a maximum of a conventional dual-view screening dose at 25 projections.

Figure 6 shows variation of the AUC values with increasing number of angular projections that span different angular arcs in the 3.6°−44.8° range using (a) the weighted average of test statistics technique and (b) the Bayesian decision fusion technique. In this case, isoimage-dose condition with only one dose fraction of 1/25 was used. At each angular range, the AUC values increase with the increase in the number of angular projections before reaching an asymptote. The number of projections at which the AUC values peak depends on the angular span. The highest AUC is obtained at an angular span of 44.8°, with 11 projections using Bayesian decision fusion technique and 17 projections with the weighted averaging of test statistics.

Figure 7 shows variation of AUC under isostudy-dose conditions using (a) the weighted average of test statistics technique and (b) the Bayesian decision fusion technique. For all angular spans, the AUC first increases and then decreases as the number of projections is increased. In this case too, the number of projections at which the AUC values peak depends on the angular span. While the maximum AUC

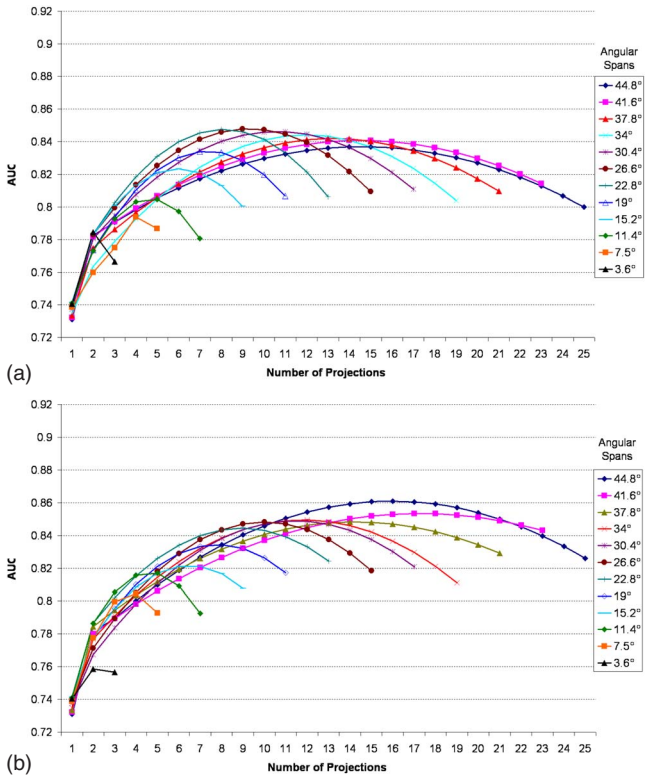


FIG. 7. Variation of AUC under isostudy-dose conditions using the (a) weighted averaging of the test statistics and (b) Bayesian decision fusion technique. The total dose level, equal to 1/25th of the standard dual-view mammographic screening dose level, was linearly divided among the different projections and, hence, the total dose delivered remains constant at this dose level.

value is obtained at an angular span of 44.8° with 16 projections using Bayesian decision fusion technique, the maximum is obtained at just 22.8° with eight projections using the weighted averaging of test statistics technique.

IV. DISCUSSION

The multiprojection imaging technique offers an advantage over single projection imaging techniques in rendering pathology that may be surrounded by a complex cloud of anatomical structures. This is accomplished by exploiting the differences in geometrical perspectives that different projections in a multiprojection scheme offer. Geometrical configuration of the data acquisition therefore plays a pivotal role in the output of a multiprojection system. Image quality in tomosynthesis, for example, can be compromised by an unoptimized data acquisition scheme. In recent implementations of tomosynthesis for breast imaging, while the total dose delivered to the patient has varied between 0.75 and 1 times the standard procedure, the number of angular projections has varied between 11 and 25 with total angular span in the 30°–60° range.[3,10,29] Because of the flexibility in the ways the images may be captured in a multiprojection system, it is important that an optimum configuration of data acquisition components be selected to realize the full potential of such a system.

The methodology used in this study quantifies the effect of changing the geometry of acquisitions in order to maximize the diagnostic image quality of a multiprojection breast CI system. A noteworthy observation from this analysis was that the two fusion techniques, namely the weighted averaging of test statistics and the Bayesian decision fusion techniques, clearly show improvement in diagnosis when information from multiple images are combined in a multiprojection CI system, indicating potentially improved breast cancer detection using a multiprojection CI system. Since the final clinical decision is based on a collective decision made from detectability cue gathered from each angular projection, it is essential to fuse the 25 ROCs into one final index of performance. The fusion techniques are, therefore, a critical element of such a multiple acquisition setup.

The two fusion approaches buildup on the framework of mathematical observer models that have been shown to correlate with detection performance of human observers on complex anatomical backgrounds. As a result, the Bayesian decision fusion and the weighted averaging of test statistics techniques potentially emulate the decision process used by human observers in arriving at the final decision. However, there are certain differences between the two fusion techniques; while the Bayesian technique draws cues from a training data set to fuse binary detection decisions from each of the 25 angular projections, the second technique explicitly incorporates the effect of correlation between the oblique angle projections and the central projection. While a similar Bayesian decision technique has been reported earlier in the literature,[27] the weighted averaging of the test statistics is a novel approach for decision fusion. As such, the two approaches show comparable effects of changing the acquisition scheme on the performance of CI and are expected to correlate with the base line detection performance of a clinician using a multiprojection system.

The ROC fusion techniques developed in this study may be extended to optimize a digital breast tomosynthesis system. An optimization scheme for tomosynthesis would, however, also incorporate the effects of reconstruction inherent in the technique. While this study establishes relative diagnostic performance of a multiprojection acquisition scheme, the findings may not be directly extrapolated to tomosynthesis without taking reconstruction into account. However, in the absence of reconstruction, the present results can be viewed as a reflection of the inherent information content of a multiprojection method.

Comparing CI to tomosynthesis, the absence of the need to reconstruct images thereby avoiding reconstruction artifacts provides an inherent advantage to CI. In addition, in contrast to tomosynthesis in which as many as 50–80 slices may need to be reviewed depending on the size of the breast, a radiologist may be presented with significantly fewer images in CI, potentially improving confidence or even accuracy of a radiologist's decision. We hypothesized that the use of a smaller number of projections in CI compared to that in tomosynthesis may be adequate to sufficiently reduce the influence of anatomical noise to yield superior diagnostic information. This conclusion is further supported by recent

studies that indicate potential improvement in diagnosis by fusing two views in stereo-mammography.[30,31] Our own study clearly shows a notable advantage of using multiple views as compared to one-view mammography at constant total dose (Fig. 7).

As currently implemented, the fusion techniques show that that the diagnostic performance of CI, in terms of detectability of an embedded mass, is dependent on both the number of projections and the total angular span of those projections. As shown in Fig. 6, the AUC first increases with an increase in the number of projections and then appears to reach an asymptote, irrespective of the angular span. This may be because as the number of projections increase, the observer increasingly incorporates information about the anatomical variation in the image until the observer has "learned" sufficient information. Beyond this level of detection, the performance becomes only quantum noise limited and, hence, dose-dependent rather than anatomical noise or projection dependent. Furthermore, the performance at the same number of projections improves with increase in the angular spans. A possible cause of this behavior is that larger angular spans likely provide diagnostic information that may not otherwise be available from smaller angular spans. This pattern, however, reverses as the number of projections decreases: the performance at smaller angular spans is better for a lower number of projections. This may be due to two reasons; first, there is higher correlation between neighboring images spanning a smaller arc than those which are spread along a larger arc, and second, at smaller angles, the x-ray beam has smaller paths resulting in lower attenuation of the beam than encountered at larger angles. This behavior is also seen in Fig. 7 that shows the variation in AUC values under isostudy-dose condition: the AUC values are higher for lower number of projections at a smaller angular span. The most noteworthy observation, however, is the sharp drop in the AUC values with increase in the number of projections, resulting in a bell-shaped characteristic of the AUC values as a function of number of projections. Besides a redundancy in anatomical information with increase in the number of projections, the sharp drop may also be attributed to the fact that although the total dose remains constant in the isostudy-dose condition, the dose level of each projection decreases with the increase in the number of projections, thereby decreasing the detectability due to reduction of the quantum signal-to-noise-ratio at each projection.

The dose dependence of detectability is clear in Fig. 5, which plots the variation of AUC as a function of number of projections for a fixed angular span of 44.8°, but different dose levels. The AUC values are seen to reach an asymptote with an increase in the number of projections, albeit the absolute value of AUC decreases, as expected, with a decrease in the dose level. Most noteworthy is the fact that the inflection point where the AUC maximizes is seen to be different for different dose levels. As a result, although a total of 11 or 17 projections at 44.8° may be deemed an optimized geometry as indicated by Fig. 6, this optimization is dose dependent. This is because the total dose delivered at any projection, under the isoimage-dose condition considered for Figs. 5 and 6, is higher than that delivered at a smaller number of projections. An alternative, however, is the isostudy-dose condition evaluated in this work in which the total dose delivered was constant. Such an evaluation is especially important in light of the observation from Fig. 6 that the performance at smaller angular spans may be better with a small number of projections. A higher total dose level at this number of projections than that possible in the isoimage-dose condition may therefore potentially increase the performance beyond that achieved at 11 or 17 projections with a total angular span of 44.8°. Incidentally, a recent study on the implementation of CI for chest imaging also found a similar number of images to render optimum performance.[11]

One limitation of this study was our inability to evaluate performance at an arbitrary number of angular projections for each angular span. This is because we were limited by the sampling of the original clinical images which was uniform within each angular span. This is evident in Figs. 6 and 7 where the AUC values were plotted for fewer angular projections for narrower angular spans. Furthermore, we recognize that in the first decision fusion approach based on a weighted average of test statistics, the weights could be defined as a function of other alternatives such as the angular separation, the total number of projections, and a reference projection other than the central projection used for this study. The technique, however, was developed as a first approximation to the actual clinical decision process used by radiologists. In addition, although the observer models have been shown to correlate well with human detection performance on real anatomical backgrounds like mammograms, an exhaustive study has not yet been carried out to conclusively establish the correlation. Furthermore, because we could only simulate a reduction of exposure levels by adding corresponding noise onto the images, only one isostudy-dose condition of the total dose of D_θ could be investigated (D_θ corresponding to 1/25 of standard two-view mammographic screening dose). The isostudy-dose trends thus reported in this study may vary at dose levels close to the clinical dose level. Finally, the optimum acquisition parameters determined for the multiprojection system in this study may not be directly applicable to tomosynthesis. Nonetheless, the methodologies developed in this study may serve as guidelines for optimizing the acquisition parameters and dose for any multiprojection imaging system.

V. CONCLUSIONS

In this study, we developed a framework to optimize the geometry of acquisitions of a multiprojection CI system by combining information from its multiple projections. It was found that the detectability of an embedded mass increased by fusing information from multiple projections demonstrating the feasibility of CI as a potential technique for improved breast cancer detection. The overall performance of the multiprojection system was a function of the number of projections used, the total angular span of those projections, and the acquisition dose level. A key finding in the situation in which the total patient dose increases with the number of

projections was that the detectability approached an asymptote at about 11–17 projections spread over an angular arc of ~45°. In situations in which the total patient dose was kept constant independent of the number of projections, the performance approximately followed a bell curve with the best detectability obtained with only eight projections spanning an angular arc of ~23°. The methodology presented here for optimizing acquisition parameters are generic in nature and may be easily adopted for optimizing the acquisition parameters for other multi-imaging techniques.

ACKNOWLEDGMENTS

The authors would like to acknowledge Craig Abbey at UC Davis for his feedback on the observer model workup. Thanks are also due to Jonathan Jesneck for helpful discussions on the Bayesian decision fusion methodology. Thanks to Swatee Singh for collecting and organizing the tomosynthesis data as a part of the ongoing clinical trial at Duke. This work was supported in part by the DOD (Grant No. W81XWH-06-1-0449), NIH (Grant Nos. R01-CA109074 and R01-CA112437), Komen Foundation (Grant No. PDF55806), and Cancer Research Prevention Foundation.

[a]Telephone: 919-684-1466; Fax: 919-684-1492. Electronic mail: asc14@duke.edu

[1]E. Samei, M. J. Flynn, and W. R. Eyler, "Relative influence of quantum and anatomical noise on the detection of subtle lung nodules in chest radiographs," Radiology **213**, 727–734 (1999).

[2]F. O. Bochud, F. R. Verdun, C. Hessler, and J.-F. Valley, "Detectability of radiological images: The influence of anatomical noise," Proc. SPIE **2436**, 156–164 (1995).

[3]X. Gong, S. J. Glick, B. Liu, A. A. Vedula, and S. Thacker, "A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam CT breast imaging," Med. Phys. **33**, 1041–1052 (2006).

[4]T. Wu, R. H. Moore, and D. B. Kopans, "Voting strategy for artifact reduction in digital breast tomosynthesis," Med. Phys. **33**, 2461–2471 (2006).

[5]L. Zhou, J. Oldan, P. Fisher, and G. Gindi, "Low-contrast lesion detection in tomosynthetic breast imaging using a realistic breast phantom," Proc. SPIE **6142**, 61425A (2006).

[6]T. Wu, R. H. Moore, and D. B. Kopans, "Voting strategy for artifact reduction in digital breast tomosynthesis," Med. Phys. **33**, 2461 (2006).

[7]E. Samei, D. M. Catarious, A. H. Baydush, C. E. Floyd, and R. Vargas-Voracek, "Bi-plane correlation imaging for improved detection of lung nodules," Proc. SPIE **5030**, 284–297 (2003).

[8]E. Samei, S. A. Stebbins, J. T. Dobbins, and J. Y. Lo, "Multiprojection correlation imaging for improved detection of pulmonary nodules," AJR Am. J. Roentgenol. **188**, 1239–1245 (2007).

[9]R. S. Saunders, E. Samei, N. Majdi-Nasab, and J. Y. Lo, "Initial human subject results for breast bi-plane correlation imaging technique," Proc. SPIE **6514**, 1–7 (2007).

[10]I. Reiser *et al.*, "Computerized mass detection for digital breast tomosynthesis directly from the projection images," Med. Phys. **33**, 482–491 (2006).

[11]E. Samei, S. A. Stebbins, J. T. Dobbins, and J. Y. Lo, "Multiprojection correlation imaging for improved detection of pulmonary nodules," AJR Am. J. Roentgenol. **188**, 1239–1245 (2007).

[12]S. Suryanarayanan, A. Karellas, S. Vedantham, H. Ved, and C. J. D'Orsi, "Detection in compressed digital mammograms using numerical observers," Proc. SPIE **5034**, 513–521 (2003).

[13]J. Yang and C. D. Cerosaletti, "Observer models and human visual detection performance with different targets," Proc. SPIE **5372**, 319–329 (2004).

[14]A. E. Burgess, "Visual signal detection with two-component noise: Low-pass spectrum effects," J. Opt. Soc. Am. A **16**, 694–704 (1999).

[15]A. E. Burgess, F. L. Jacobson, and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise," Med. Phys. **28**, 419–437 (2001).

[16]A. S. Chawla, E. Samei, R. Saunders, C. Abbey, and D. Delong, "Effect of dose reduction on the detection of mammographic lesions: A mathematical observer model analysis," Med. Phys. **34**, 3385–3398 (2007).

[17]R. S. Saunders, E. Samei, and C. Hoeschen, "Impact of resolution and noise characteristics of digital radiographic detectors on the detectability of lung nodules," Med. Phys. **31**, 1603–1613 (2004).

[18]E. Samei, M. Flynn, and W. Eyler, "Simulation of subtle lung nodules in projection chest radiography," Radiology **202**, 117–124 (1997).

[19]L. T. Niklason *et al.*, "Digital tomosynthesis in breast imaging," Radiology **205**, 399–406 (1997).

[20]J. T. Dobbins and D. J. Godfrey, "Digital x-ray tomosynthesis: Current state of the art and clinical potential," Phys. Med. Biol. **48**, R65–R106 (2003).

[21]I. Sechopoulos, S. Suryanarayanan, S. Vedantham, C. J. D'Orsi, and A. Karellas, "Scatter radiation in digital tomosynthesis of the breast," Med. Phys. **34**, 564–576 (2007).

[22]R. S. Saunders and E. Samei, "A method for modifying the image quality parameters of digital radiographic images," Med. Phys. **30**, 3006–3017 (2003).

[23]M. P. Eckstein, C. K. Abbey, and E. O. Bochud, "The effect of image compression in model and human performance," Proc. SPIE **3663**, 243–252 (1999).

[24]M. P. Eckstein, C. K. Abbey, and J. S. Whiting, "Human vs. model observers in anatomic backgrounds," Proc. SPIE **3340**, 16–26 (1998).

[25]Y. Zhang, B. Pham, and M. Eckstein, "Task-based model/human observer evaluation of SPIHT wavelet compression with human visual system-based quantization," Acad. Radiol. **12**, 324–336 (2005).

[26]H. H. Barrett, C. K. Abbey, and B. Gallas, "Stabilized estimates of Hotelling observer detection performance in patient-structured noise," Proc. SPIE **3340**, 27–43 (1998).

[27]J. L. Jesneck, L. W. Nolte, J. A. Baker, C. E. Floyd, and J. Y. Lo, "Optimized approach to decision fusion of heterogeneous data for breast cancer diagnosis," Med. Phys. **33**, 2945–2954 (2006).

[28]I. Sechopoulos, S. Suryanarayanan, S. Vedantham, C. J. D'Orsi, and A. Karellas, "Computation of the glandular radiation dose in digital tomosynthesis of the breast," Med. Phys. **34**, 221–232 (2007).

[29]Y. Zhang *et al.*, "A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis," Med. Phys. **33**, 3781 (2006).

[30]A. D. A. Maidment, P. R. Bakic, and M. Albert, "Effects of quantum noise and binocular summation on dose requirements in stereoradiography," Med. Phys. **30**, 3061–3071 (2003).

[31]D. J. Getty and P. J. Green, "Clinical applications for stereoscopic 3-D displays," J. Soc. Inf. Displ. **23**, 36 (2007).

Effect of dose reduction on the detection of mammographic lesions: A mathematical observer model analysis

Amarpreet S. Chawla
Duke Advanced Imaging Laboratories, Departments of Radiology and Biomedical Engineering, Duke University Medical Center, Durham, North Carolina 27705

Ehsan Samei and Robert Saunders
Duke Advanced Imaging Laboratories, Departments of Radiology, Physics and Biomedical Engineering, Duke University Medical Center, Durham, North Carolina 27705

Craig Abbey
Department of Biomedical Engineering, University of California, Davis, California 95616

David Delong
Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina 27705

(Received 2 June 2006; revised 6 June 2007; accepted for publication 15 June 2007; published 30 July 2007)

The effect of reduction in dose levels normally used in mammographic screening procedures on the detection of breast lesions were analyzed. Four types of breast lesions were simulated and inserted into clinically-acquired digital mammograms. Dose reduction by 50% and 75% of the original clinically-relevant exposure levels were simulated by adding corresponding simulated noise into the original mammograms. The mammograms were converted into luminance values corresponding to those displayed on a clinical soft-copy display station and subsequently analyzed by Laguerre-Gauss and Gabor channelized Hotelling observer models for differences in detectability performance with reduction in radiation dose. Performance was measured under a signal known exactly but variable detection task paradigm in terms of receiver operating characteristics (ROC) curves and area under the ROC curves. The results suggested that luminance mapping of digital mammograms affects performance of model observers. Reduction in dose levels by 50% lowered the detectability of masses with borderline statistical significance. Dose reduction did not have a statistically significant effect on detection of microcalcifications. The model results indicate that there is room for optimization of dose level in mammographic screening procedures. © *2007 American Association of Physicists in Medicine*. [DOI: 10.1118/1.2756607]

Key words: dose reduction, simulation, mammography, Laguerre-Gauss, Gabor, hotelling observer model, AUC, ROC, SKEV

I. INTRODUCTION

Early detection of breast cancer is largely contingent on the image quality of mammograms. There are many aspects of image quality. Among those, it highly depends on the amount of x-ray exposure and, thus, the patient dose, used to form the images. Existing guidelines define upper limits of 4 mGy for the mean glandular doses (MGD) at a representative breast thickness[1] with clinical values ranging between 0.3–12 mGy.[2] However, currently there are no strict protocols for optimizing radiation dose in mammography such that dose may be minimized without compromising diagnostic quality.

Since the breast is a highly radio-sensitive organ, it is important that the dose level during screening mammography be kept as low as possible so that screening in itself does not unduly increase the risk of breast cancer induction. Research studies have aimed to optimize the benefit-to-radiation risk ratio of digital mammographic screening procedures.[3] Berns *et al.* showed that dose values can be lower for digital systems than screen-film systems.[4] A recent study by Gennaro *et al.*[5] on phantoms concluded that this dose reduction factor could be as much as 50% while still preserving image quality above screen/film standards. Besides reducing the adverse effects of radiation dose, another rationale for lowering dose of x-ray mammography comes from the fact that detectability of lesions in mammograms is confounded by overlapping anatomical structures—a characteristic inherent to most projection imaging techniques. In fact, it has been suggested that anatomical variability could be the biggest factor limiting the detection of breast cancer, both by radiologists and by computer assisted diagnosis (CAD) systems.[6,7] This limitation of mammography is not governed by the amount of x-ray exposure and, hence, the dose level delivered to the breast. Thus, it is contended that dose levels in mammographic procedures can potentially be reduced with a lesser impact on the detectability of lesions than that caused by anatomical variations.

One way to assess the impact of reduced dose levels on the detection of breast lesions is to use mathematical model observers. Such observers have been used in the past to predict the performance of human observer in clinically realistic tasks.[8–12] Most previous studies based on such observers

have used raw (unprocessed digital) images as direct input to the observer models, thereby overlooking the fact that, in a clinical situation, the digital image values are converted to luminance values for the radiologist's interpretation. To emulate realistic diagnostic tasks, it is only logical that luminance values corresponding to the digital values of mammograms be used as input to observer models, though rarely implemented.[13] Furthermore, a majority of previous studies have only used a signal-known-exactly (SKE) detection paradigm where the same lesion whose shape and size are known *a priori* is embedded on variable backgrounds which are either real[8,9] or computer generated.[11,12] Recently, signal-known-exactly-but-variable (SKEV) framework was introduced as a better approximation to clinically relevant task of a signal known as statistically paradigm.[14,15] Under SKEV paradigm, the shape and size of a signal are changed from one image to another. Prior implementations of SKEV have used a forced choice paradigm without a complete receiver operating characteristics (ROC) analysis.[10]

In this work, mathematical model observers were used to investigate the impact of reduced dose on the detection of breast lesions. The detectability of simulated masses and microcalcifications at reduced dose levels were analyzed and compared with the detectability on mammograms acquired with clinical dose levels. To simulate a clinically relevant situation, first raw (unprocessed digital) images obtained from a mammography system were processed with techniques typical of clinical mammograms. Next, the post-processed images were mapped to luminance values before using them as inputs to model observers. Furthermore, using a SKEV diagnostic paradigm with variable backgrounds, a complete ROC analysis was performed and the area under the ROC curves (AUC) computed to compare detectability at different dose levels.

II. MATERIALS AND METHODS

A. Image database

Images used in this study were clinically acquired at 25–30 kVp with molybdenum anode using a clinical flat-panel cesium iodide-based digital mammography system (Senographe 2000D, GE Medical Systems, Milwaukee, WI). Based on normal radiologist reports, a total of 300 normal craniocaudal view mammograms were pooled. From each of these mammograms, a region of interest (ROI) of size 512×512 (5.12×5.12 cms taking into account the detector pixel size of 100 μm) was extracted. The resultant 300 ROIs were used as the signal (lesion)-absent set of backgrounds for training the observer models. Another set of signal-present images were generated by digitally inserting realistic simulated masses and microcalcifications in signal-absent backgrounds by a routine previously published.[16,17] This routine relied on the measured characteristics of real lesions to create simulated lesions with a realistic appearance. Four different types of lesions were produced: typically malignant masses (modeled after irregular ill-defined and irregular spiculated masses), typically benign masses (modeled after oval circumscribed and oval obscured masses), and typically malignant microcalcifications (fine linear branching and clustered pleomorphic microcalcifications). The lesion sizes chosen led to masses with a diameter of 3.3–4.1 mm while individual microcalcifications of 0.35 mm in average diameter were located inside microcalcification distributions with 4–7 mm in diam. The lesion contrast was set assuming a heterogeneous breast (50% glandular/50% adipose tissue, representing an average breast composition) and accounting for the applicable target/filteration combination, acquisition kVp, detector material, and appropriate scatter fractions.

Fifty different realizations of each of the four lesion types were simulated generating, in total, 200 lesion realizations. They were added to the mammograms in a logarithmic scale to model the x-ray attenuation process. A lesion was placed at the center of a mammographic background, creating a total of 50 signal-present images for each lesion type. The same 50 backgrounds were used for each lesion type. To simulate real clinical situations, the backgrounds corresponding to the 50 signal-present images were different from the 250 backgrounds without signal. Figure 1 shows examples of the simulated lesions.

Following the insertion of the lesions, a noise modification routine was used to add radiographic noise to clinically-acquired mammograms to create images with a noise appearance similar to that caused by a reduction in radiation dose. Noise equivalent of a particular dose level was determined using a novel algorithm, which has been reported earlier in detail.[18] The routine accounted for the quantum noise variance, the detector transfer properties, the scattered radiation, and the impact of varying attenuation due to breast structures. By changing the noise magnitude, dose reduction corresponding to half (50%) and quarter (25%) of the original clinical exposure levels were simulated. Image post-processing techniques typical of clinical mammograms were applied to the images with the appropriate window and level parameters for each mammogram determined by an experienced mammographer. Figure 2(a) shows an ROI of a typical mammogram with a malignant mass located at the center of the image. Figures 2(b) and 2(c) show the same mammogram with added simulated noise corresponding to one half and a quarter of the full-dose level, respectively.

B. Luminance transformation

To simulate a realistic clinical setup, each of the ROIs from step A were mapped into luminance values according to the expected display luminance transformation/calibration function. The specific display function used in this study was the DICOM GSDF which has been recommended by the TG18 committee and is used as the *de facto* standard for grayscale calibration of soft-copy displays.[19] The images were transformed to luminance domain via accordance with an eight-bit look-up table covering DICOM calibration between a minimum luminance of 0.5 cd/m^2 and a maximum luminance of 300 cd/m^2. The luminance-mapped ROIs were then used as input to the observer models.

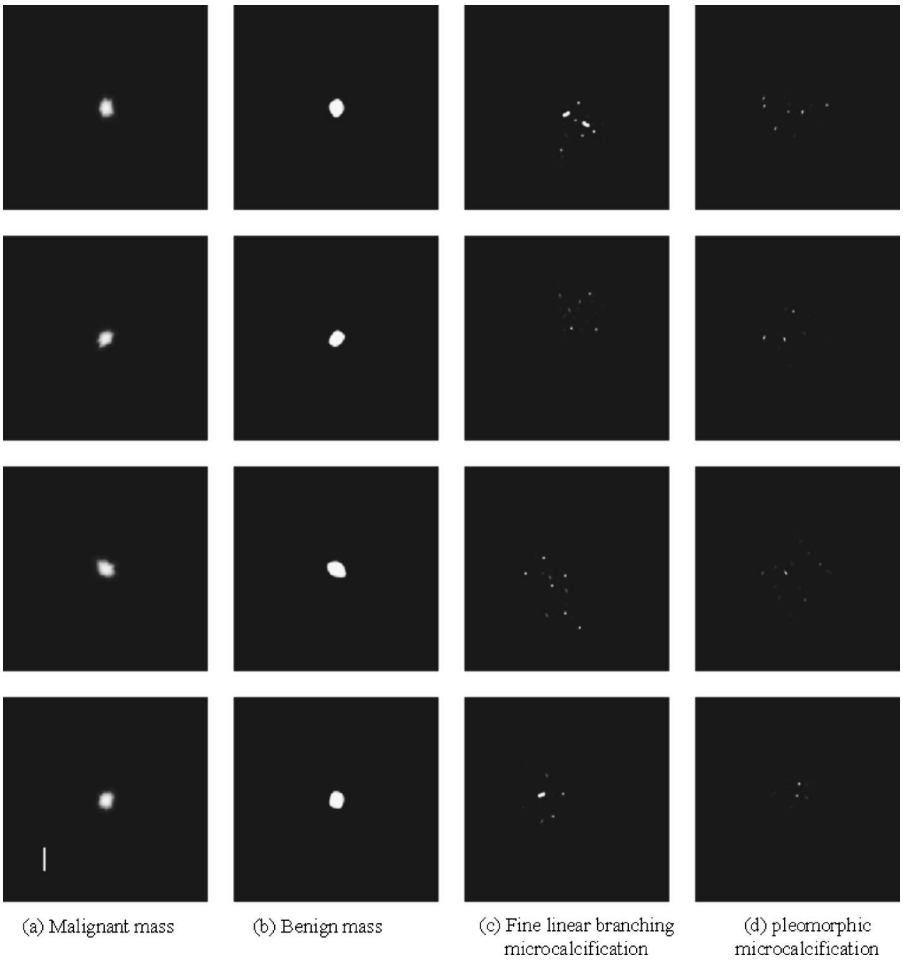


FIG. 1. Example of four different realizations of simulated lesions: (a) malignant mass, (b) benign mass, (c) fine linear branching microcalcifications, and (d) pleomorphic (pleo) malignant microcalcifications. Benign masses and fine line branching microcalcifications were also simulated, but have not been shown here. These lesions were embedded on real mammographic backgrounds.

C. Observer models

Observers models have been used to predict human observer performance in clinically relevant visual tasks such as the detection of lesions in a complex background structure.[8–10] In this study we used a specific implementation of a particular class of observer models, the Hotelling observer. Hotelling observer is a linear observer. It has been shown earlier that linear observers are good predictors of human visual performance under various tasks involving simulated pathology embedded in real image backgrounds.[8–10,20–22] Hotelling observer uses a linear function of the pixels of an image and takes correlation in the image background into account to compute its test statistics. Consequently, it requires an estimate of the sample covariance matrix from the original images to incorporate statistical fluctuations in the image. That estimation from the limited number of mammograms clinically available necessitates a reduction in the size of the sample covariance matrix. Towards that end, computationally simpler implementation of Hotelling observers, called Channelized Hotelling Observer (CHOs), are used.[23] In this study, two such functions were used, namely the Laguerre-Gauss functions, which model an ideal observer, and an anthropomorphic set of channels, called Gabor functions, which assess the visual system of human observer. These functions extract essential linear features from the image, thereby considerably reducing the dimensionality of the problem.

1. Laguerre-Gauss channelized hotelling observer (LG CHO)

The Laguerre-Gauss channels are smooth functions which are a product of Laguerre polynomials and Gaussians. The functions use a distance scale[8] related to the signal radius which defines the variance of the Gaussian. It is generally iteratively adjusted to maximize the area under the ROC curves. For the lesion size in the 3–7 mm range used in this study, a value of 8 for the distance scale was found to maximize the AUC.

A pilot experiment was conducted to investigate the number of channels required to satisfactorily represent the lesion template in the channel space. Toward this end, backgrounds with four levels of white Gaussian noise with standard deviations of 0.25, 0.5, 1, and 2 were generated. Simulated masses identical to those used in the main study (Fig. 1) were inserted on these backgrounds to generate signal-present images. A simple template-matching ideal observer was constructed and its performance, in terms of AUC, was compared with that of LG CHO. An average performance of

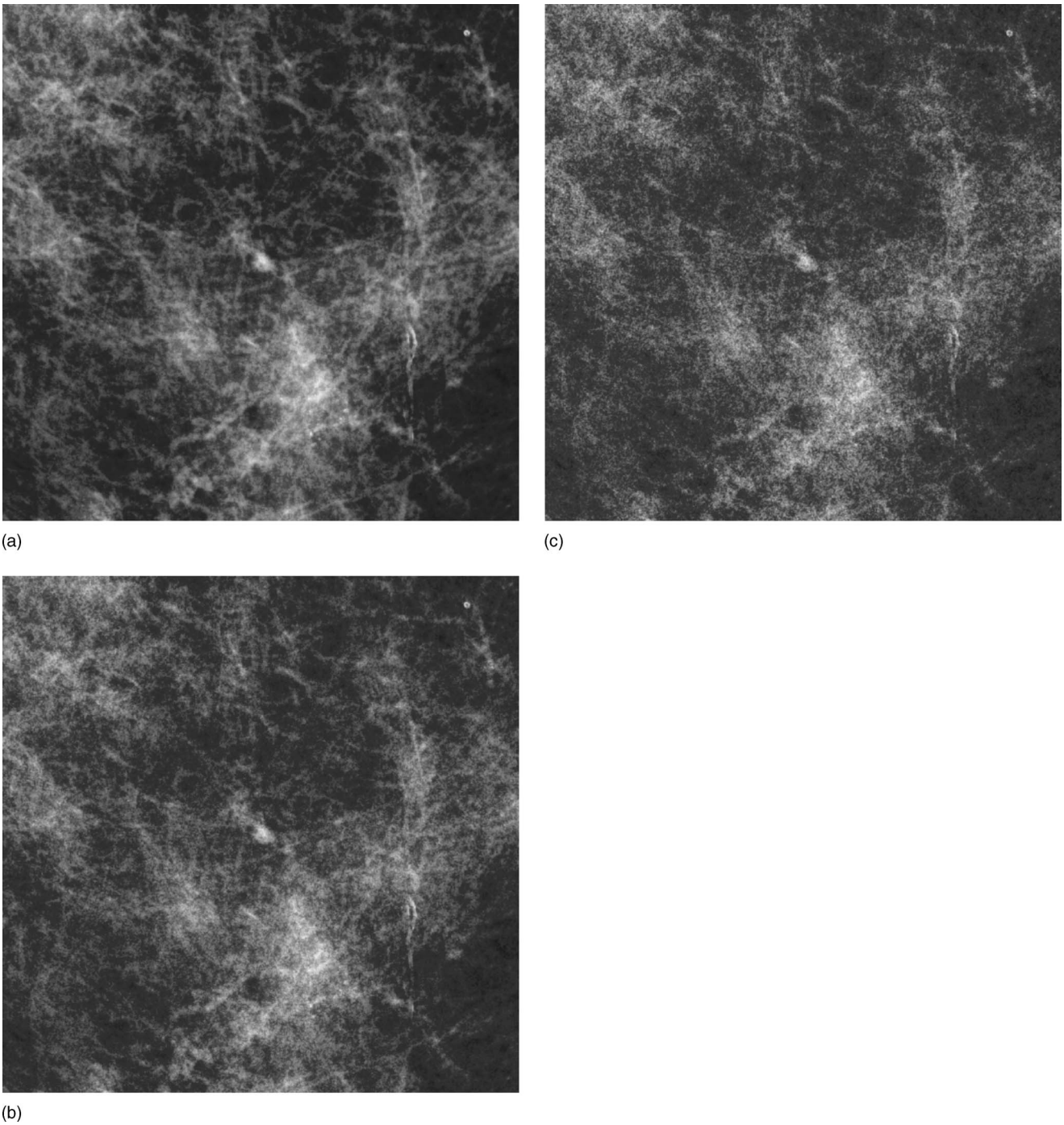


FIG. 2. Example ROIs of signal-present mammograms with a malignant lesion present at the center. (a) shows an ROI of a mammogram acquired at the clinical dose level. (b) and (c) show the same ROI with added noise corresponding to half and quarter of the full-dose level, respectively.

the two observers was computed, in line with the SKEV paradigm followed throughout in this study, with different realizations of the noise at each of the four noise levels of the background. It was found that a total of ten channels provides a stable representation of the lesion template. To confirm these findings, another experiment was conducted using the mammographic backgrounds and the four lesion types used for the main study. In this case the performance of LG CHO was measured in terms of AUCs as a function of number of channels. A total of ten channels was found to maximize the observer performance. Observer performance reaches an asymptote with further increase in the number of channels. These results indicated that a total of ten LG channels is sufficient to represent the four lesion types, even though generally six channels are considered adequate for characterizing isotropic signals.[23]

2. Gabor channelized Hotelling observer

While LG channels are efficient in modeling the ideal observer in tasks involving Gaussian noise and Gaussian

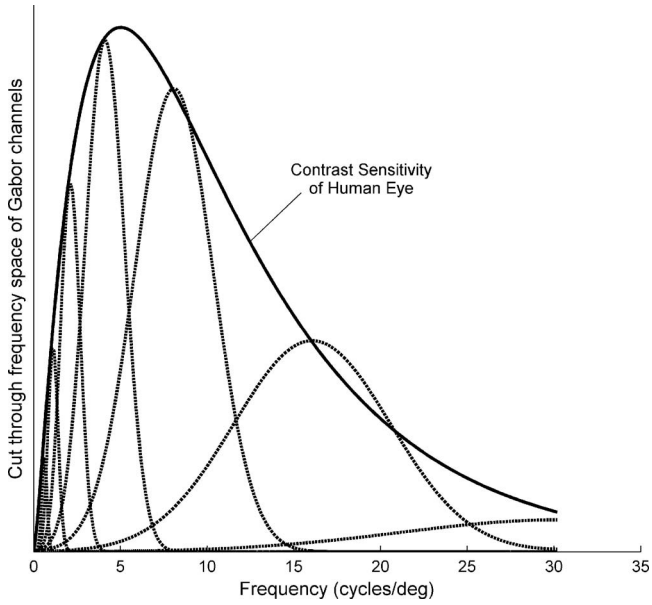


FIG. 3. Comparison of frequency response of Gabor channels with contrast sensitivity of the human eye. Octave bandwidth of Gabor channels was 1, centered at frequencies (0.5,1,2,4,16,32) cycles/degree, each oriented at 0°.

backgrounds, they do not mimic the contrast sensitivity of the human eye. With widespread use in physiological and psychophysical experiments, Gabor functions are based on the response of neurons to a small spot of light as a function of position.[24] Figure 3 shows how Gabor channels sample the contrast sensitivity of the human eye. The figure also illustrates the band-pass characteristics of the Gabor channels. Per earlier psychophysical studies,[11,25] in our implementation, the channels had an octave bandwidth of 1 and orientations of 0°, 45°, 90°, and 135° to sample the full range of orientational dependencies. Central frequencies chosen were 0.5, 1, 2, 4, 8, 16, 32 cycles/degree. These, with the odd and even phases add up to a total of 56 channels. The lower limit of central frequencies, 0.5 cycles/deg, was chosen to incorporate the lesions (whose size varied from 4–8 mm) and a sufficient surrounding background of about 110 mm around the lesions. This keeps the context information intact around the lesions and also corresponds to the region to which the human eye gets adapted to when looking at an image from a typical viewing distance of 50 cm.[26] The upper limit of the chosen central frequencies was the octave frequency closest to the contrast sensitivity limit of the human eye, which is at 30 cycles/degree.[27] This can be seen in Fig. 3 which shows that one of the channels has maximum gain near 30 cycles/degree. For a typical soft-copy display pixel size of 0.144 mm, 30 cycles/degree is also close to the Nyquist frequency of 30.30 cycles/deg at a typical viewing distance of 50 cm.

Having determined the parameters of the two model observers, they were next applied to the images. The channel responses were calculated as the dot-product between the channels vector and the lesion (signal)-present image vector, lesion (signal)-absent image vector (also referred to as background vector), and lesion (signal)-only image vector. The channel weights corresponding to each ith lesion were determined as

$$w_c^i = (K_c)^{-1} R_s^i, \quad i = 1, \ldots, 50, \tag{1}$$

where K_c is the covariance matrix of the channel responses of the mammographic backgrounds and R_s^i is the channel response to ith lesion-only image. The covariance matrix was determined from the signal-absent images since the inserted lesions were of low contrast and did not affect the covariance matrix.

Finally, the decision variables for signal-present and signal-absent cases were computed as the dot-product of the channel weights and channel responses of the signal-present and signal-absent mammograms, respectively. For each ith signal-present image, decision variable corresponding to signal-present case was computed as dot-product of that image with the ith lesion template, whereas decision variables for signal-absent cases were calculated as dot-products of signal-absent backgrounds with the ith lesion template. Thus, for each lesion i, a single decision variable corresponding to signal-present case was computed, whereas 250 decision variables were computed for the 250 signal-absent backgrounds available in the database. This was repeated for all the 50 lesion templates available in the databases. These decision variables were thus computed by

$$\lambda_1^i = (w_c^i)^t R_{g1}^i \quad \text{for } i\text{th lesion present image},$$

$$\lambda_0^{ij} = (w_c^i)^t R_{g0}^j \quad \text{for } j\text{th lesion-absent background}, \tag{2}$$

where R_{g1}^i is the response of an ith mammographic background with ith lesion embedded in it, and R_{g0}^j is the response of jth signal-absent background; i varies between 1 and 50, and j varies from 1 to 250. This methodology is in line with the SKEV paradigm in which each of the 50 signal-present images had a different but known lesion present. Fifty ROC curves were, thus, obtained for the 50 lesions using this procedure. These curves were non-parametric in nature and were derived by simple thresholding on the probability density function (pdf) of the decision variables. The ROC curves were then averaged along their true-positive fractions to generate a final mean ROC. The area under the mean ROC curve (AUC) was determined using the trapezoidal rule. Finally, detectability index, d_A, were computed from the AUC values as $2\,\mathrm{erf}^{-1}(2\mathrm{AUC}-1)$, where erf^{-1} is the inverse of error function. A schematic of the methodology is shown in Fig. 4.

D. Evaluation of statistical significance

Statistical resampling methods (i.e., bootstrap) were used to test the statistical significance of differences in AUCs obtained for the detection at different dose levels.[28] This was performed as follows: At each dose-level, 500 bootstrap samples of 50 signal-present images and of 250 background images were generated. For each bootstrap sample, an ROC curve and AUC were determined by the method described earlier. This procedure was repeated for all 500 bootstrap

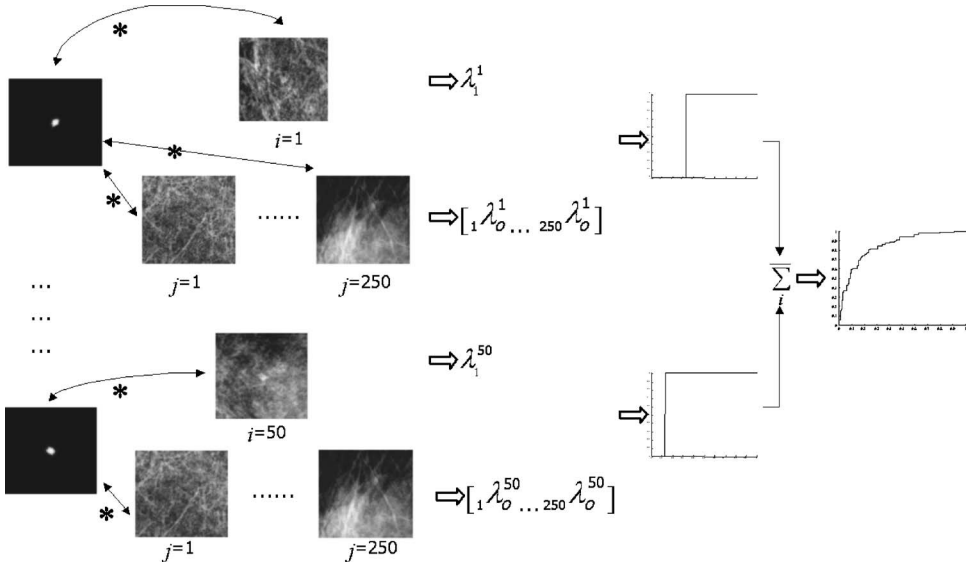


FIG. 4. Schematic for the computation of ROC curve for signal know exactly but variable (SKEV) task. The image database had 50 different lesion-present images (denoted with subscript *i*), each with a different lesion embedded in it. There were 250 signal-absent or "normal" mammographic backgrounds (denoted with subscript *j*). The symbol "*" denotes dot-product.

samples obtaining a set of 500 ROCs and, thus, a set of 500 AUCs for each dose-level. These sets of 500 AUC values were then subtracted in pair and the mean of differences and the standard deviation of the differences were computed as an approximation to the standard error of the difference of the means of the AUCs. Differences in AUCs were compared using one-sided *z*-tests assuming that the dose reduction would only decrease the detectability performance. If the ratio of the mean of differences and the standard error of differences, denoted as the *z*-value, was more than 1.64, the differences seen in AUCs at different dose levels were denoted as being statistically significant with a 95% confidence interval.

III. RESULTS

Figures 5 and 6 show mean ROC curves obtained from Laguerre-Gauss (LG) and Gabor channelized Hotelling observers (CHOs), respectively. The corresponding AUCs with their associated standard deviations are plotted in Fig. 7 and tabulated in Table I. These AUC values were consistent with prior literature.[29–35] The sensitivity at 70% (and 90%) specificity for the three dose levels obtained by LG CHO were 100% (95%), 100% (95%), and 98% (90%) for benign masses; 100% (92%), 98% (90%), and 98% (90%) for malignant masses; 67% (50%), 79% (52%) and 76% (45%) for fine linear branching microcalcifications; and 72% (30%), 68% (28%), and 64% (22%) for pleomorphic microcalcifications. The sensitivity at 70% (and 90%) specificity level for the three dose levels obtained by Gabor CHO were 100% (95%), 95% (83%), and 88% (72%) for benign masses; 100% (96%), 100% (95%), and 98% (88%) for malignant masses; 55% (35%), 72% (35%), and 62% (35%) for fine linear branching microcalcifications; and 60% (45%), 68% (45%), and 55% (32%) for pleomorphic microcalcifications.

The mean of pair-wise differences in AUCs obtained from 500 bootstrap samples of the decision variables at each dose levels are plotted in Fig. 8. Also plotted are the standard errors. These values are listed in Table I. A paired one-sided *z*-test evaluation of the results obtained from Gabor CHO showed, with a 95% confidence interval, that the differences in AUCs for the detection of malignant masses and microcalcifications with a 50% dose reduction were not statistically significant (*z*-values <1.64). However, this reduction in dose causes statistically significant differences in detectability of benign masses. LG CHO indicates significant differences in the detectability of benign masses as well as pleomorphic microcalcification, with a 50% reduction in dose. With reduction in dose level by 75%, Gabor CHO indicates significant difference in detectability of benign and malignant masses, whereas LG CHO indicates significant differences in the detection of benign masses and pleomorphic microcalcifications. Regardless of statistical significance, all comparisons indicate a drop in detectability of masses with a reduction in dose.

Figure 9 shows AUCs and the standard errors obtained when the model calculation was done with the digital values of the images. These values may be compared with those shown in Fig. 7 that were obtained on luminance-mapped images.

IV. DISCUSSION

In this study, we analyzed the effect of reduction in dose levels normally used in mammographic screening procedures on the detection of four breast lesion types. Towards that end, two mathematical observer models were implemented. Gabor CHO and LG CHO algorithmic observers suggested that a dose reduction by 50% only modestly influences the detection probability of malignant masses and fine linear branching microcalcifications on mammograms. A statistically significant difference was observed for this dose reduction in the detection of benign lesions by both LG and Gabor CHO and for pleomorphic microcalcification by LG CHO. However, the statistical significance of the differences in these cases are in the borderline judging from the *z*-values which are close to 2, the critical value for statistical significance in a two-sided *z*-test. The ROC curve for the detection

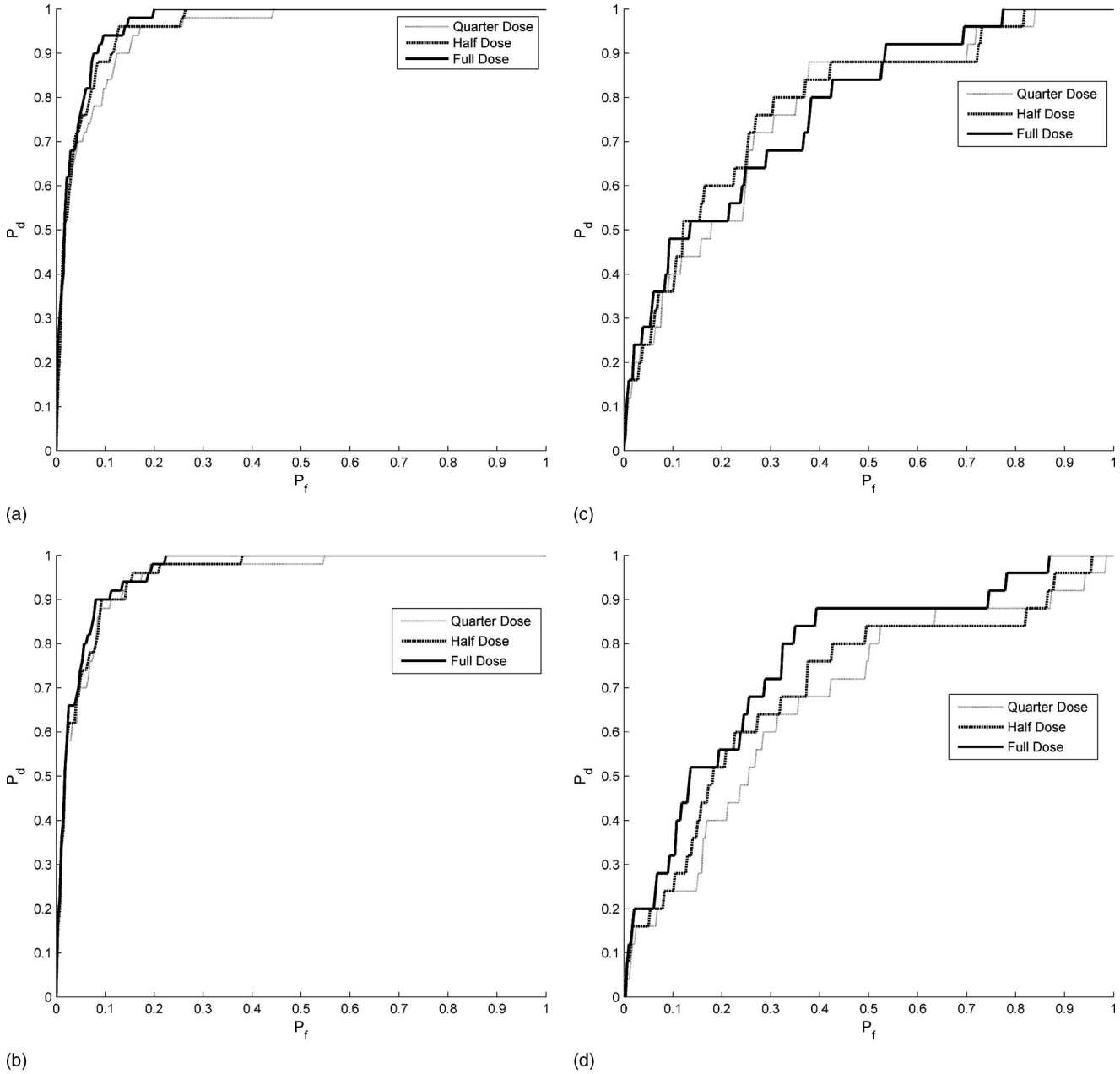


FIG. 5. ROC curves obtained from Laguerre–Gauss channelized Hotelling observer as a measure of detectability of four different lesion types at three dose levels. The four lesion types were (a) benign and (b) malignant masses, (c) fine linear branching and (d) pleomorphic microcalcifications. Full-dose denotes the typical clinical dose level used in standard mammographic screening procedures.

of benign masses with Gabor CHO at the full-dose level is significantly higher at the low sensitivity portion of the ROC curves compared to that obtained at 50% dose level as compared to ROC curves obtained from LG CHO. For reduction of dose by 75%, there is an indication, as expected, that the performance may deteriorate more than when the dose is reduced by half. Statistically significant differences in the detectability of benign masses and pleomorphic microcalcifications are observed in the performance of LG CHO and in the detectability of benign and malignant masses in the performance of Gabor CHOs when the dose is reduced by 75%.

A reduction in microcalcification detectability due to dose reduction in digital mammography was reported in a previous study.[36] However, the study was performed on anthropomorphic phantom. The results in this study indicate that reduction in dose level by as much as 50% does not significantly impact detectability of microcalcifications. Based on the range of dose levels considered in this study, it may be concluded that, in contrast to screen/film mammography in which a recent trend has been to increase radiation dose,[37] any further increase in clinical dose already delivered may not improve the detectability of mammographic lesions in digital mammography. Furthermore, although not statistically significant, absolute means of differences in AUCs are higher for the case of microcalcifications than for masses,

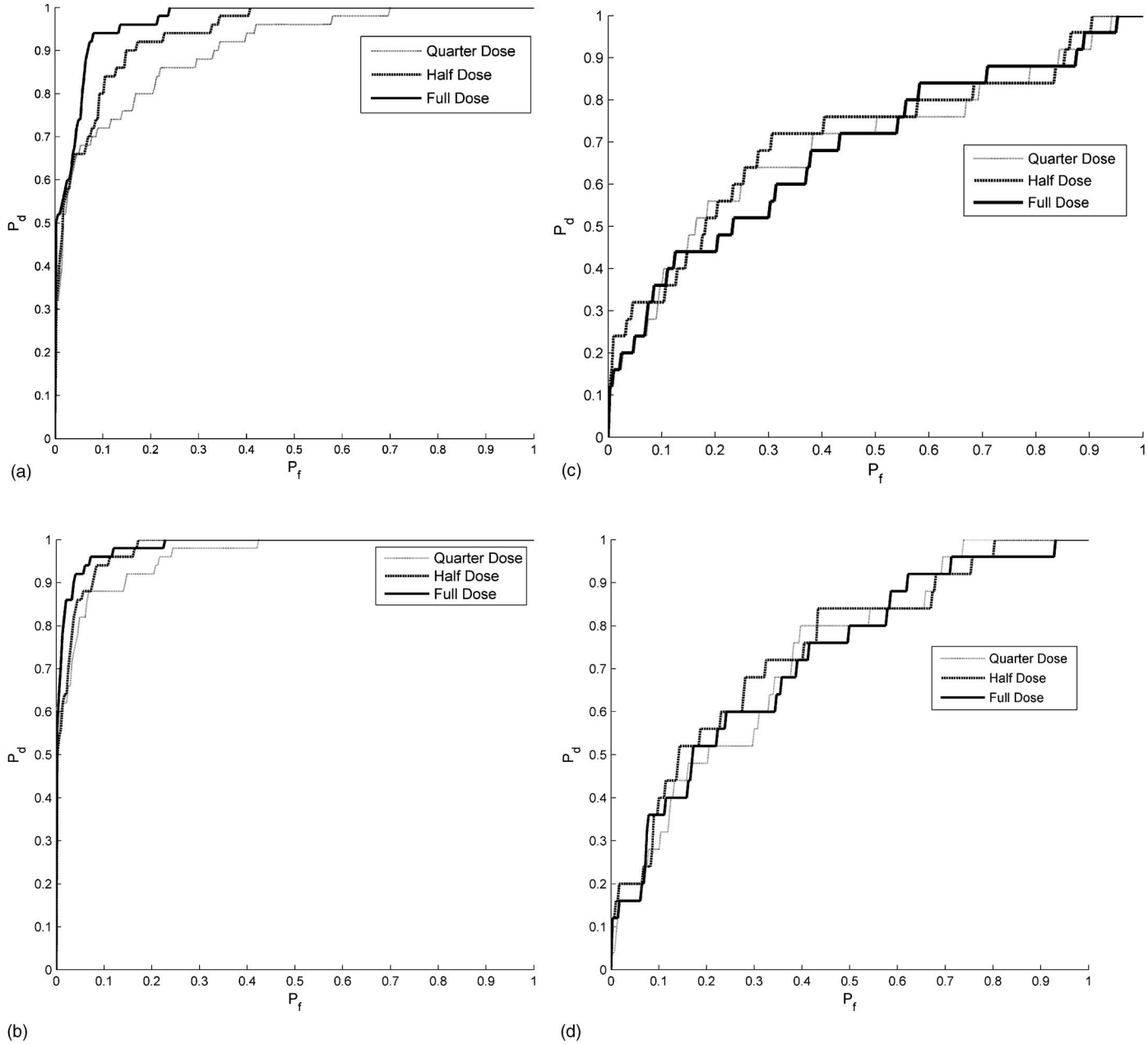


FIG. 6. ROC curves obtained from 56 Gabor channelized Hotelling observer as a measure of detectability of four different lesion types at three dose levels. The four lesion types were (a) benign and (b) malignant masses, (c) fine linear branching and (d) pleomorphic microcalcifications. Full-dose denotes the typical clinical dose level used in standard mammographic screening procedures.

indicating that a reduction of dose may affect the detection of microcalcifications more than that of masses.

In real clinical practice, a radiologist interprets digital mammograms as displayed and limited by the output of a soft-copy display. The digital values on mammograms are converted into display luminance values using a nonlinear lookup table which leads to changes in image contrast and also quantization errors since current displays are not capable of displaying images at the inherent bit-depth of the acquired images. These may have a significant effect on the detectability of breast lesions. In order to emulate that clinical situation in this study, the digital mammographic ROIs were transformed into equivalent luminance values similar to actual clinical practice. The algorithmic observers were then applied to these luminance-transformed images. To the best of our knowledge, such an analysis has rarely been reported in the literature, where analyses are often performed only on the pixel values.

To gauge the difference in detectability performance that may occur due to luminance transformation, we applied the algorithmic observers to both luminance-mapped mammographic images and to digital images themselves. The AUCs and the mean of pair-wise differences in AUCs for luminance-mapped images are tabulated in Table I and plotted in Fig. 7, while those for digital mammographic ROIs are plotted in Fig. 9. Comparing the AUCs, we observe that

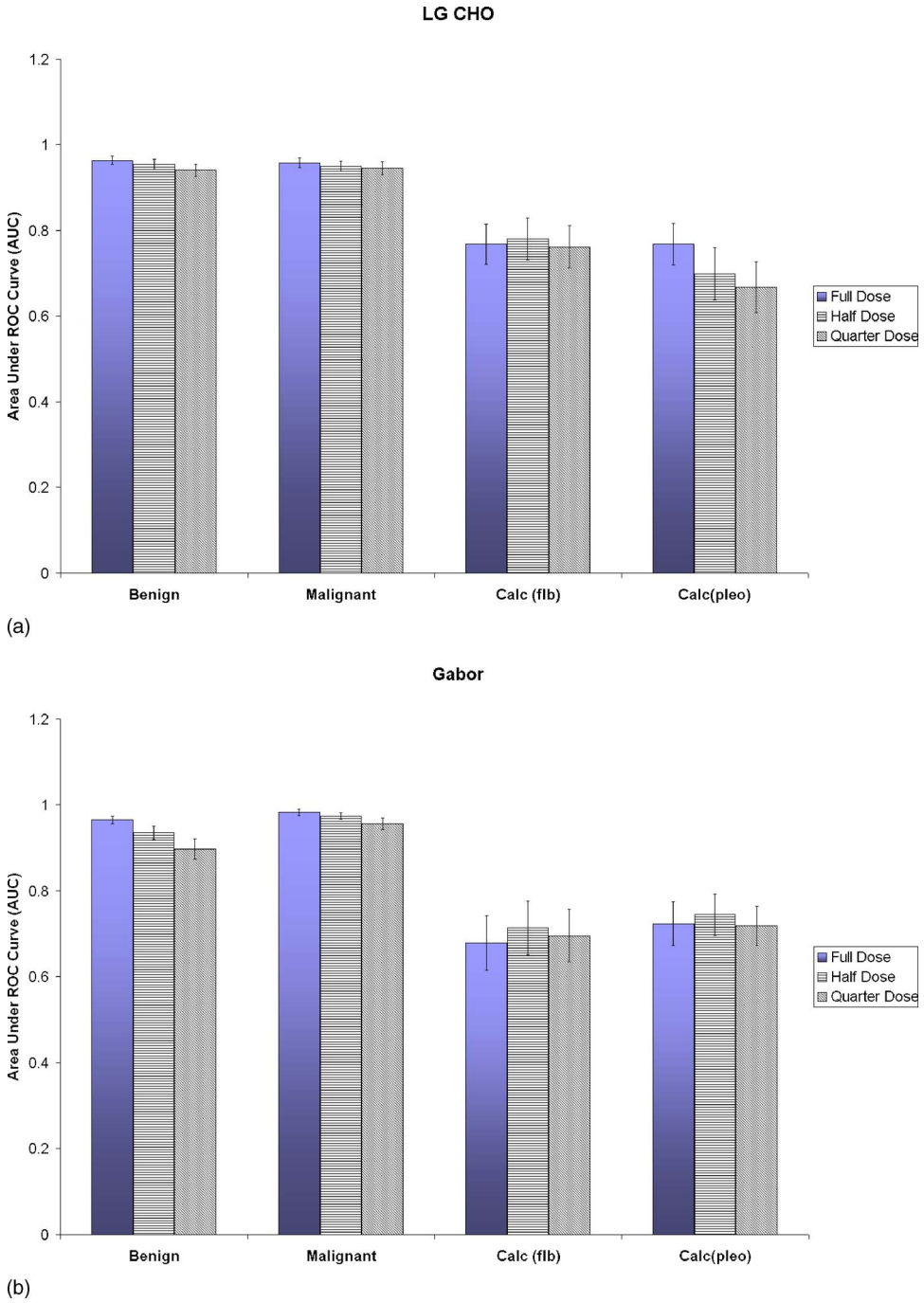


FIG. 7. Illustration of computed area under the ROC curves (AUC) and the standard deviation observed in the measurement of AUCs. AUCs are used to quantify detection performance on mammograms acquired at three different dose levels. They were obtained from the two channelized Hotelling observers, namely, (a) Laguerre–Gauss and (b) Gabor. Full-dose denotes the clinical dose level used in standard mammographic screening.

these values decrease for luminance-transformed images for all lesion-types, indicating reduction in performance as mammograms go through transformation from their digital values to equivalent luminance levels on a soft-copy display.

To further investigate the effect of luminance transformation on the performance of Hotelling observer, we conducted a pilot experiment in which the raw images (which had a 12-bit resolution) were converted into eight-bit images. The conversion was linear, i.e. no non-linear transformation was applied as is done during luminance transformation. However, the conversion to eight-bits accounts for the quantization effect that eight-bit luminance transformation of raw images causes. We found that the AUC values for eight-bit images were close to those obtained on 12-bit raw images, implying that quantization has minimal effect on detection and that the difference in detection between luminance-mapped mammographic images and digital images may largely be attributed to the nonlinear transformation inherent to luminance mapping. Although application of a nonlinear Hotelling observer, instead of the linear observer used in this study, could have inverted the effect of luminance transformation back to linear domain to improve the performance of the luminance transformed images, such a nonlinear Hotelling observer is difficult to validate and, hence, was not used in this study. Within the constraints of our observer model construct, therefore, we found that luminance transformation

TABLE I. AUCs (± std. dev.), ΔAUC (*z*-stat value) obtained with (a) LG CHO and (b) Gabor CHO on luminance-mapped mammograms. These figures of merit are a measure of differences in detection performance at full, half, and quarter dose levels. AUCs were obtained from 500 bootstrap samples of the decision variables. *z*-stat values are a measure of significance of the null hypothesis that there is no deterioration in performance with reduction in dose. The critical value for one-sided *z*-test is 1.64 at 5%.

(a)

Figure of merit	LG CHO							
	d_A (±std.dev.)			AUC (±std.dev.)			ΔAUC (*z*-stat value)	
	Full	Half	Quarter	Full	Half	Quarter	Full-half	Full-quarter
Benign mass	2.602	2.464	2.301	0.964	0.954	0.941	0.009	0.023
	±0.174	±0.172	±0.181	±0.009	±0.011	±0.014	(1.77)	(2.42)
Malignant mass	2.505	2.392	2.322	0.958	0.950	0.946	0.008	0.01
	±0.180	±0.181	±0.196	±0.010	±0.012	±0.015	(1.18)	(1.19)
Calc (flb)	1.059	1.111	1.043	0.768	0.780	0.762	−0.012	0.006
	±0.226	±0.242	±0.231	±0.047	±0.049	±0.049	(0.53)	(0.23)
Calc (Pleo)	1.024	0.757	0.641	0.768	0.698	0.667	0.069	0.10
	±0.214	±0.235	±0.221	±0.048	±0.060	±0.059	(1.75)	(2.25)

(b)

Figure of merit	Gabor CHO							
	d_A (±std.dev.)			AUC (±std.dev.)			ΔAUC (*z*-stat value)	
	Full	Half	Quarter	Full	Half	Quarter	Full-half	Full-quarter
Benign mass	2.626	2.202	1.832	0.965	0.935	0.897	0.03	0.07
	±0.184	±0.181	±0.199	±0.008	±0.015	±0.024	(2.46)	(3.19)
Malignant mass	3.067	2.817	2.481	0.983	0.975	0.956	0.008	0.027
	±0.238	±0.186	±0.204	±0.007	±0.007	±0.01	(1.36)	(2.14)
Calc (flb)	0.681	0.785	0.711	0.679	0.714	0.695	0.035	0.017
	±0.247	±0.264	±0.260	±0.063	±0.063	±0.062	(0.744)	(0.33)
Calc (Pleo)	0.845	0.942	0.862	0.724	0.745	0.719	0.021	0.004
	±0.224	±0.223	±0.211	±0.050	±0.048	±0.045	(0.66)	(0.1059)

of raw images reduces detectability and that the level of reduction is task dependent. Thus we highly recommend that all observer model calculations be performed on images converted into luminance space. It may be noted, however, that there are other aspects of display properties, such as noise power spectrum (NPS), the modulation transfer function (MTF), veiling glare, uniformity, etc., which might also affect the appearance of medical images in an actual clinical setup. Evaluation of these aspects of image quality is, however, beyond the scope of this study.

CHOs have been used in the past to predict human observer performance on images with real anatomical backgrounds, including mammograms. However, the optimum number of channels required in CHO for the best representation of images in the reduced transformation space remains an open question. The number of Gabor channels used in previous studies have varied from 16 to 80,[11,20,38,39] although there is psychophysical evidence that the visual system may be using 16 channels.[40] In the present study, 56 channels were used. There has also been a large variation in the literature in number of Laguerre-Gauss channels used. Whereas generally six channels are sufficient for characterizing isotropic signals,[23] the number of channels have varied from 6 to 18.[9,39] Burgess has suggested that as many as 40 channels may be required for characterizing nonisotropic signals.[8]

We investigated the optimum number of Laguerre-Gauss channels required for a stable observer model performance in detecting the particular lesion types used in this study. Toward this end, AUC values were determined as a function of the number of channels for detection of masses embedded on white Gaussian noise backgrounds and compared with AUC values obtained by ideal observer on the same backgrounds. Since the Hotelling observer is equivalent to an ideal observer in this case, the standard results of ideal observer can be compared against the results of the analysis by LG CHO. The results indicated that ten Laguerre-Gauss channels pro-

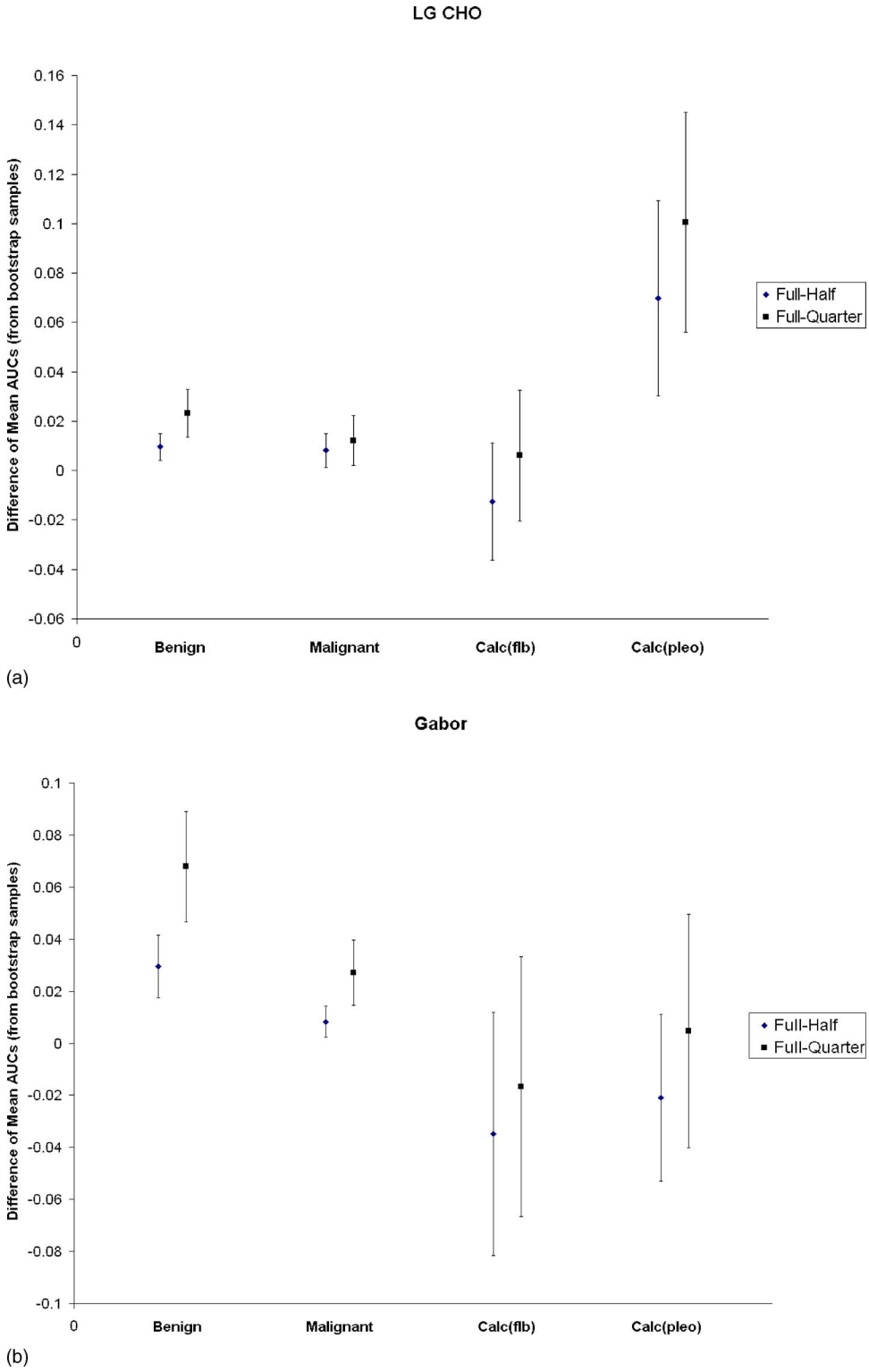


FIG. 8. Paired *z*-test analysis: Difference of mean AUCs obtained from (a) Laguerre–Gauss and (b) Gabor CHOs. The AUCs were obtained by bootstrap sampling of decision variables at different dose levels. Standard errors are also shown.

vide AUC values very close to that of an ideal observer. There are, however, two limitations of this experiment. One, since a symmetry of lesion shape is required to establish a correlation between Hotelling observer and an ideal observer, microcalcifications could not be used because they are found in clusters and lack the symmetry found otherwise in masses. Second, clinical mammograms could not be used for this analysis since the results of an ideal observer may be traced on white stationary noise backgrounds.

To confirm if ten Laguerre-Gauss channels are sufficient for clinical detection tasks, a second experiment was conducted on mammographic backgrounds embedded with the four lesion types used in the main study. AUC values were determined as a function of the number of channels. This

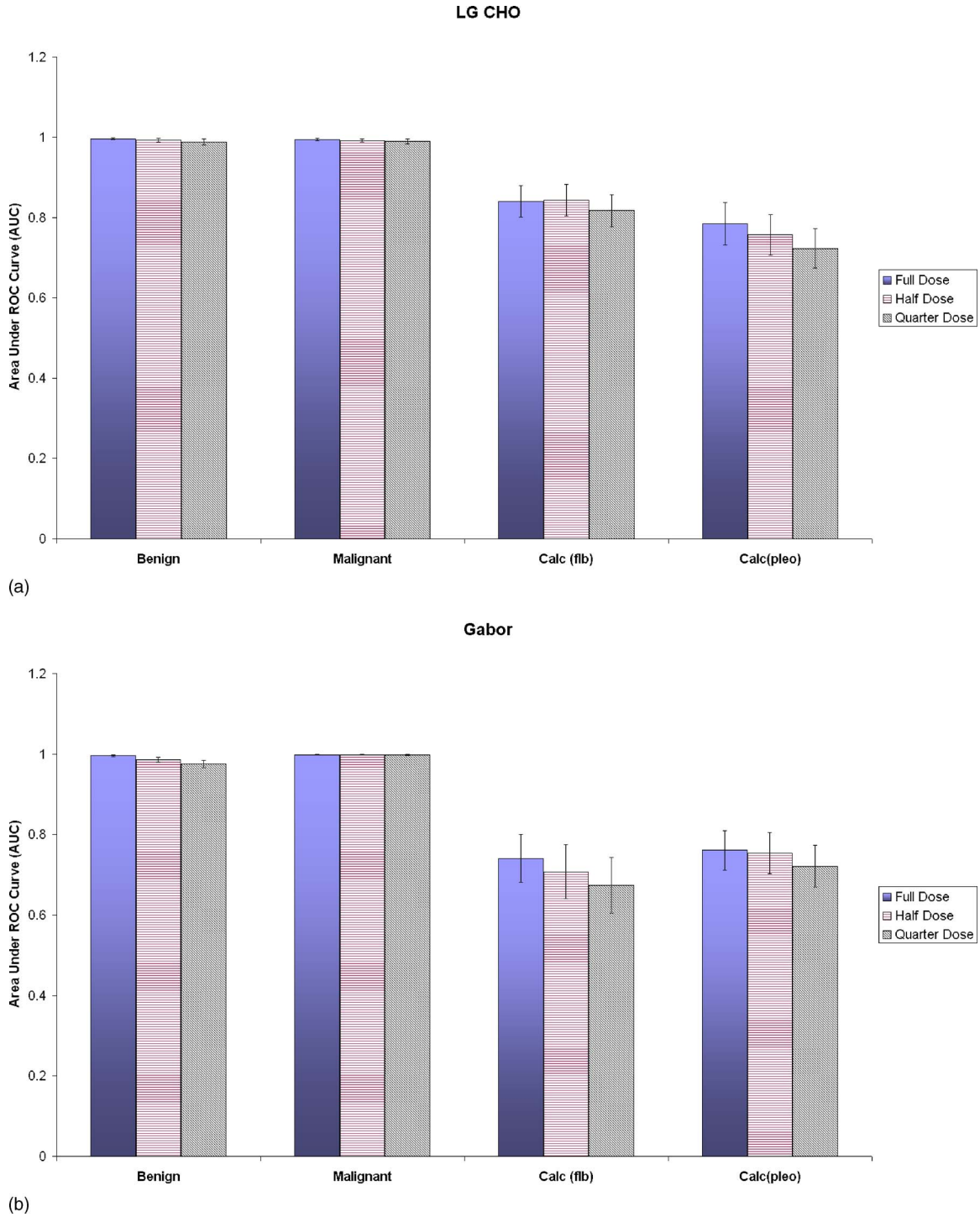


FIG. 9. Illustration of computed area under the ROC curves (AUC) and the standard deviation observed in the measurement of AUCs at three different dose levels. They were obtained from the two channelized Hotelling observers, namely, (a) Laguerre-Gauss and (b) Gabor. These values were obtained on digital mammograms as against those shown in Fig. 7 that were obtained on luminance-mapped mammograms.

variation is shown in Fig. 10. In the case of masses, no appreciable change in the AUC values was observed with increase in the number of channels. However, in the case of microcalcifications, the AUC values increased with the increase in the number of channels, and then seem to approach an asymptote. This change appears to occur at ten channels. Based on this finding and also that from the first experiment, we used a total of 10 LG channels in this study. Moreover, the use of ten channels results in a covariance matrix of size 10×10 and, hence, the 300 mammographic backgrounds available for training were sufficient to provide a stable estimate of the population statistics.

The difference in the number of channels used may be one reason why the detection performances of LG CHO and

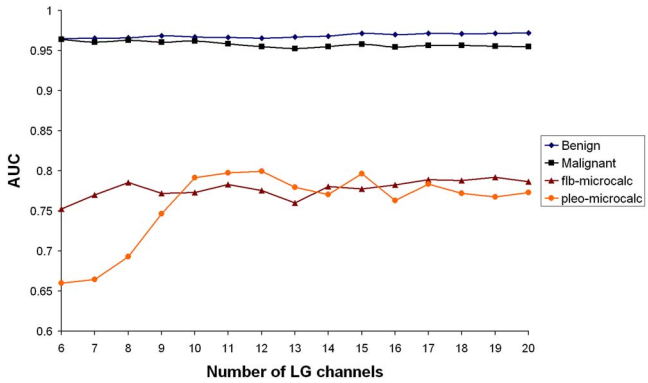


FIG. 10. Change in AUC values obtained from LG CHO as a function of number of channels used. This analysis was used to determine the number of LG channels required for a stable estimate of the lesion template for the four lesion types. Analysis was done on mammographic backgrounds used in the main study.

Gabor CHO differ, as is clear from comparing Fig. 7(a) and 7(b). These differences may also be attributed to the departure of LG channels' response from that of Gabor channels. In the frequency response of Gabor channels, there are regions which are more weighted than others, as well as regions which are not covered at all. Moreover, Gabor channels, unlike LG channels, are not isotropic and therefore do not assume isotropicity of masses. Furthermore, while the Gabor channels extended upto the Nyquist frequency of a typical soft-copy display, the frequency coverage of the ten LG channels used in this study was limited to about half of the Nyquist frequency.

The d_A values reported in this study provide us with a performance metric but not the metric for which they are commonly used. This is because the test statistics used to compute these values may not be Gaussian and so they can only be interpreted as a result of some function which is monotonically related to AUC values. Finally, it was not possible to report the standard error of the difference of d_A obtained from the bootstrap samples as has been done for AUC values. This is because for many such bootstrapped samples, AUC values obtained were 1, consequently making the d_A values infinity, thus rendering the calculation of standard error cumbersome. Therefore, only standard deviations of d_A obtained from the 50 signal-present images have been reported.

Notwithstanding the conclusions of this investigation, certain limitations should be acknowledged. One limitation of this research was the low number of images used for testing the performance of the mathematical observer models. The resulting limited statistical significance along with the small magnitude of performance reduction with dose generally conform with the findings of another research group[5] which concluded that a dose reduction by as much as 50% from the currently practiced clinical dose levels may not compromise clinical decisions. It may be noted that these conclusions are based on the assumption that the differences in detectability observed are independent of the specific values of AUCs obtained in this study.

Furthermore, our study was based on specific conspicuity levels of the simulated lesions. Although increasing the conspicuity level would have improved detection, especially that of microcalcifications, the current level was chosen to be consistent with the follow-up human observer study.[41] This was done based on an assumption that the findings would translate to different (higher and lower) conspicuity levels seen in clinical situations. This is a necessary assumption for any study of this type. At the same time, AUC values noted in this work are consistent with those previously reported in the literature.[29–35] In addition, although the observer models have been shown to correlate well with human detection performance on real anatomical backgrounds, like mammograms, an exhaustive study has not yet been carried out to conclusively establish the correlation. Moreover, LG CHO model requires symmetry (isotropy) of the signal for robust template computation, which microcalcifications, in particular, clearly lack. This might affect the sensitivity values for their detection with LG CHO. At the same time, the performance of LG CHO was close to and, hence, substantiated by that of Gabor CHO model that does not require the signal to be symmetric. Notwithstanding, it is important that the significant potential of 50% dose reduction found in this study should be implemented only after being confirmed clinically. Finally, an important issue of the effect of reduced dose on discrimination between benign and malignant masses has not been addressed here.

V. CONCLUSIONS

In summary, patient dose in mammographic screening should be optimized based on clinical decision (that is the final outcome of the radiological process). Different observer models used in this research show that reduction in dose-level by 50% did affect detectability of masses although with borderline statistical significance. Dose reduction by 50%, however, did not have a statistically significant reduction in detection of microcalcifications except in one case where the reduction was marginally significant.

More importantly, the observer models indicate that there is a potential for the reduction of dose level in mammographic screening procedures without severely compromising the detectability of masses. Based on the range of dose levels considered in this study, the results also suggest that any increase in dose level than the clinical levels currently used may not improve diagnostic performance. The present findings need additional confirmation by rigorous clinical trials and human observer studies before being implemented clinically.

[1]National Council on Radiation Protection and Measurements (NCRP), Mammography–A user's guide, Report 85, Bethesda, MD, 1986.
[2]R. L. Kruger and B. A. Schueler, "A survey of clinical factors and patient dose in mammography," Med. Phys. **28**, 1449–1454 (2001).
[3]J. Law and K. Faulkner, "Concerning the relationship between benefit and radiation risk, and cancer detected and induced, in breast screening programme," Br. J. Radiol. **75**, 678–684 (2002).
[4]E. A. Berns, R. E. Hendrick, and G. R. Cutter, "Performance comparison of full-field digital mammography to screen-film mammography in clinical practice," Med. Phys. **29**, 830–834 (2002).

[5]G. Gennaro, L. Katz, H. Souchay, C. Alberelli, and C. diMaggio, "Are phantoms useful for predicting the potential of dose reduction in full-field digital mammography?" Phys. Med. Biol. **50**, 1851–1870 (2005).
[6]F. O. Bochud, F. R. Verdun, C. Hessler, and J.-F. Valley, "Detectability of radiological images: the influence of anatomical noise," Proc. SPIE **2436**, 156–164 (1995).
[7]R. M. Nishikawa, C. J. Vyborny, M. L. Giger, and K. Doi, "Analysis of false-positive microcalcification clusters identified by a mammographic computer-aided detection scheme," Proc. SPIE **2167**, 773–777 (1994).
[8]A. E. Burgess, F. L. Jacobson, and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise," Med. Phys. **28**, 419–437 (2001).
[9]S. Suryanarayanan, A. Karellas, S. Vedantham, H. Ved, and C. J. D'Orsi, "Detection in compressed digital mammograms using numerical observers," Proc. SPIE **5034**, 513–521 (2003).
[10]Y. Zhang, B. Pham, and M. Eckstein, "Task-based model/human observer evaluation of SPIHT wavelet compression with human visual system-based quantization," Acad. Radiol. **12**, 324–336 (2005).
[11]J. Yang and C. D. Cerosaletti, "Observer models and human visual detection performance with different targets," Proc. SPIE **5372**, 319–329 (2004).
[12]A. E. Burgess, "Visual signal detection with two-component noise: low-pass spectrum effects," J. Opt. Soc. Am. A **16**, 694–704 (1999).
[13]A. Badano and B. D. Gallas, "Detectability decreases with off-normal viewing in medical liquid crystal displays," Acad. Radiol. **13**, 210–218 (2006).
[14]M. P. Eckstein and C. K. Abbey, "Model observers for signal-known-statistically tasks (SKS)," Proc. SPIE **4324**, 91–102 (2001).
[15]M. P. Eckstein, Y. Zhang, and B. Pham, "Optimization of model observer performance for signal known exactly but variable tasks leads to optimized performance in signal known statistically tasks," Proc. SPIE **5034**, 123–134 (2003).
[16]R. S. Saunders, E. Samei, J. Johnson, and J. Baker, "Effect of display resolution on the detection of mammographic lesions," Proc. SPIE **5749**, 243–250 (2005).
[17]R. Saunders, E. Samei, J. Baker, and D. Delong, "Simulation of mammographic lesions," Acad. Radiol. **13**, 860–870 (2006).
[18]R. S. Saunders and E. Samei, "A method for modifying the image quality parameters of digital radiographic images," Med. Phys. **30**, 3006–3017 (2003).
[19]E. Samei *et al.*, American Association of Physicists in Medicine (AAPM), Assessment of Display Performance for Medical Imaging Systems, Task Group 18, (Madison, WI, April 2005).
[20]M. P. Eckstein, C. K. Abbey, and E. O. Bochud, "The effect of image compression in model and human performance," Proc. SPIE **3663**, 243–252 (1999).
[21]B. D. Gallas and H. H. Barrett, "Validating the use of channels to estimate the ideal linear observer," J. Opt. Soc. Am. A **20**, 1725–1738 (2003).
[22]M. P. Eckstein, C. K. Abbey, and J. S. Whiting, "Human vs. model observers in anatomic backgrounds," Proc. SPIE **3340**, 16–26 (1998).
[23]H. H. Barrett, C. K. Abbey, and B. Gallas, "Stabilized estimates of Hotelling observer detection performance in patient-structured noise," Proc. SPIE **3340**, 27–43 (1998).
[24]D. G. Stork and H. R. Wilson, "Do Gabor functions provide appropriate descriptions of visual cortical receptive fields?" J. Opt. Soc. Am. A **7**, 1362–1373 (1990).
[25]A. B. Watson, O. J. Braddick, and A. C. Sleigh, *Detection and recognition of simple spatial forms* (Springer-Verlag, Berlin, 1983), pp. 100–114.
[26]W. L. Makous, "Fourier models and the loci of adaptation," J. Opt. Soc. Am. A **14**, 2323–2345 (1997).
[27]P. G. J. Barten, *Contrast Sensitivity of the Human Eye and its Effects on Image Quality* (SPIE, Bellingham, 1999), pp. 7–66.
[28]B. Efron and G. Gong, "A leisurely look at the bootstrap, the jackknife and cross-validation," Am. Stat. **37**, 36–48 (1983).
[29]X. J. Rong *et al.*, "Microcalcification detectability for four mammographic detectors: Flat-panel, CCD, CR, and screen/film," Med. Phys. **29**, 2052 (2002).
[30]Y. Higashida *et al.*, "Detection of subtle microcalcifications: comparison of computed radiography and screen-film mammography," Radiology **183**, 483–486 (1992).
[31]C.-J. Lai *et al.*, "Receiver operating characteristic analysis for the detection of simulated microcalcifications on mammograms using hardcopy images," Phys. Med. Biol. **51**, 3901–3919 (2006).
[32]S. Suryanarayanan, A. Karellas, S. Vedantham, S. M. Waldrop, and C. J. D'Orsi, "Detection of simulated lesions on data-compressed digital mammograms," Radiology **236**, 31–36 (2005).
[33]A. R. Cowen, J. H. Launders, M. Jadav, and D. S. Brettle, "Visibility of microcalcifications in computed and screen-film mammography," Phys. Med. Biol. **42**, 1533–1548 (1997).
[34]J. Oestmann *et al.*, "Comparison of two screen-film combinations in contact and magnification mammography: detectability of microcalcifications," Radiology **168**, 657–659 (1988).
[35]P. Skaane *et al.*, "Breast lesion detection and classification: Comparison of screen-film mammography and full-field digital mammography with soft-copy reading—Observer performance study," Radiology **237**, 37–44 (2005).
[36]S. Obenauer, K. Hermann, C. Schorn, U. Fischer, and E. Grabbe, "Full-field digital mammography: dose-dependent detectability of simulated breast lesions," Fortschr. Rontgenstr **172**, 1052–1056 (2000).
[37]E. D. Pisano *et al.*, "Factors affecting phantom scores at annual mammography facility inspections by the U.S. Food and Drug Administration," Acad. Radiol. **8**, 864–870 (2001).
[38]O. Nestares, R. Navarro, J. Portilla, and A. Tabernero, "Efficient spatial-domain implementation of a multiscale image representation based on Gabor functions," J. Electron. Imaging **7**, 166–173 (1998).
[39]Y. Zhang, B. T. Pham, and M. P. Eckstein, "Evaluation of internal noise methods for Hotelling observers," Proc. SPIE **5749**, 162–173 (2005).
[40]J. G. Daugman, "Spatial visual channels in the Fourier plane," Vision Res. **24**, 891–910 (1984).
[41]E. Samei, R. S. Saunders, J. A. Baker, and D. M. Delong, "Digital mammography: Effects of reduced radiation dose on diagnostic performance," Radiology **243**, 396–404 (2007).